PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



(51) International Patent Classification 5:		(11) International Publication Number:	WO 93/13663
A01N 43/22, A61K 31/71 C07H 17/08, C12N 1/21 C12P 19/62		(43) International Publication Date:	22 July 1993 (22.07.93)
(21) International Application Number: PCT/US (22) International Filing Date: 17 January 1992 (BE, CH, DE, DK, ES, FR, G	
(71) Applicant: ABBOTT LABORATORIES [US/US] D-377/AP6D-2, One Abbott Park Road, Abbott 60064-3500 (US).	; CHA Park,	D Published IL With international search report	t.
(72) Inventors: KATZ, Leonard; 844 North Avenue gan, IL 60085 (US). DONADIO, Stefano; 110 Road, Libertyville, IL 60048 (US). MCALPINI B.; 211 West Rockland Road, Libertyville, (US).	Brook! E, Jam	ill es,	
(74) Agents: GORMAN, Edward, Hoover, Jr. et al.; A boratories, CHAD D-377/AP6D-2, One Abb Road, Abbott Park, IL 60064-3500 (US).			

(54) Title: METHOD OF DIRECTING BIOSYNTHESIS OF SPECIFIC POLYKETIDES

(57) Abstract

A method to produce novel polyketide structures by designing and introducing specified changes in the DNA governing the synthesis of the polyketide is disclosed. The biosynthesis of specific polyketide analogs is accomplished by genetic manipulation of a polyketide-producing microorganism by isolating a polyketide biosynthetic gene-containing DNA sequence, identifying enzymatic activities associated within the DNA sequence, introducing one or more specified changes into the DNA sequence which codes for one of the enzymatic activities which results in an altered DNA sequence, introducing the altered DNA sequence into the polyketide-producing microorganism to replace the original sequence, growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog, and isolating the specific polyketide analog from the culture. The method is most useful when the segment of the chromosome modified is involved in an enzymatic activity associated with polyketide biosynthesis, particularly for manipulating polyketide synthase genes from Saccharapolyspora or Streptomyces.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

	_	50	France	MR	Mauritania
AT	Austria	FR		MW	Malawi
AU	Australia	GA	Gabon		Netherlands
BB	Barbados	GB	United Kingdom	NL	
BE	Belgium	GN	Guinca	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
	•	IE	Ireland	PT	Portugal
BJ	Bunin	IТ	Italy	RO	Romania .
BR	Brazil		•	RU	Russian Federation
ÇA	Canada	JP	Japan*	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic		Sweden
CG	Congo		of Korea	SE	
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	('ôte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Cameroon	LJ	Liechtenstein	SU	Soviet Union
	Czechoslovakia •	LK	Sri Lanka	TĐ	C:had
cs		1.U	Luxembourg	TG	Togo
CZ	Czech Republic		•	UA	Ukraine
DΕ	Germany	MC	Monaco	US	United States of America
ÐK	Denmark	MG	Madagascar		Viet Nam
ES	Spain	Ml.	Mali	VN	AIGI LIUM
CI	Einland .	MN	Moneolia		

-

15

20

25

30

35

METHOD OF DIRECTING BIOSYNTHESIS OF SPECIFIC POLYKETIDES

Field of the Invention.

The present invention relates to a method for directing the biosynthesis of specific polyketide analogs by genetic manipulation. In particuar, polyketide biosynthetic genes are manipulated to produce precise, novel polyketides of predicted structure.

10 Background of the Invention

Polyketides are a large class of natural products that includes many important antibiotics and immunosuppressants such as erythromycins, tetracyclines, and rapamycins. Their synthesis proceeds by an ordered condensation of acyl esters to generate carbon chains of varying length and substitution pattern that are later converted to mature polyketides. This process has long been recognized as resembling fatty acid biosynthesis, but with important differences. Unlike a fatty acid synthase, a typical polyketide synthase is programmed to make many choices during carbon chain assembly: For example, the choice of "starter" and "extender" units, which are often selected from acetate, propionate or butyrate residues in a defined sequence. The choice of using a full cycle of reductiondehydration-reduction after some condensation steps, omitting it completely, or using one of two incomplete cycles (reduction alone or reduction followed by dehydration), which determines the pattern of keto or hydroxyl groups and the degree of saturation at different points in the chain is additionally programed. Finally the choice of stereochemistry for the substituents at many of the carbon atoms is programmed by the polyketide synthase.

Because of the commercial significance of *Streptomyces*, a great amount of effort has been expended in the study of *Streptomyces* genetics. Consequently much is known about *Streptomyces* and several cloning vectors exist for transformations of the organism.

Although many polyketides have been identified, there remains the need to obtain novel polyketide structures with enhanced properties. Current methods of obtaining such molecules include screening of natural isolates and chemical modification of existing polyketides, both of which are costly and time consuming. Current screening methods are

based on gross properties of the molecule, i.e. antibacterial, antifungal

10

15

20

25

30

35

activity, etc., and both *a priori* knowledge of the structure of the molecules obtained or predetermination of enhanced properties are virtually impossible. Chemical modification of preexisting structures has been successfully employed, but it still suffers from practical limitations to the type of compounds obtainable, largely connected to the poor yield of multistep syntheses and available chemistry to effect modifications. The following modifications are extremely difficult or inefficient at the present time: change of the stereochemistry of the side chains in the completed polyketide; change of the length of the polyketide by removal or addition of carbon units from the interior of the acyl chain; and dehydroxylation at unique positions in the acyl chain. Accordingly, there exists the need to obtain molecules wherein such changes can be specified and performed and would represent an improvement in the technology to produce altered polyketide molecules with predicted structure.

Summary of the Invention

The present invention provides a method to produce novel structures from designing and introducing specified changes in the DNA governing the synthesis of the polyketide. According to the method of the present invention, the biosynthesis of specific polyketide analogs is accomplished by genetic manipulation of a polyketide-producing microorganism comprising the steps of:

- (1) isolating a polyketide biosynthetic gene-containing DNA sequence;
- (2) identifying enzymatic activities associated within said DNA sequence;
- (3) introducing one or more specified changes into said DNA sequence which codes for one of said enzymatic activities which results in an altered DNA sequence;
- (4) introducing said altered DNA sequence into the polyketideproducing microorganism to replace the original sequence;
- (5) growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog; and
 - (6) isolating said specific polyketide analog from the culture.

The present method is most useful when the segment of the chromosome modified is involved in an enzymatic activity associated with polyketide biosynthesis. The present invention is especially useful



10

in manipulating polyketide biosynthetic genes from *Streptomyces*, an organism which provides over one-half of the clinically useful antibiotics.

Brief Description of the Drawings

FIG. 1 illustrates the organization of gene encoding polyketide synthase and designated *eryA* as follows: (a) Map coordinates of the DNA; (b) DOTPLOT of the output of COMPARE (window = 50, stringency = 32) program (Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin, Biotechnology Center, 1710 University Avenue, Madison, Wisconsin 53705 of *eryA* segment (x-axis) vs. subsegment of *eryA* comprises between 23 - 27.5 sequence coordinates (y-axis) [see Fig. 2]; (c) Open reading frame organization of *eryA* and enzymatic activities encoded. PT = propionyltransferase; ACP = acyl carrier protein; KS= β -ketoacyl ACP synthase; RmT = (2R) methylmalonyl CoA

CoA transferase; KR =β-ketoredu mase; SmT = (2S) methylmalonyl CoA transferase; DH = dehydratase; ER = enoylreductase; TE = thioesterase; and (d) Schematic diagram showing the extent of each of the six modules in eryA.

FIG. 2. illustrates the nucleotide sequence of *eryA* with corresponding translation of the three open reading frames. Standard one letter codes for the amino acids appear beneath their respective nucleic acid codons. The standard one letter codes for the amino acid sequences are as follows:

25 Α -alanine R -arginine N -asparagine D -aspartic acid C -cysteine 30 Q -glutamine E -glutamic acid G -glycine Η -histidine Ι -isoleucine 35 L -leucine K -lysine M -methionine (start) F -phenylalanine

P -proline
S -serine
T -threonine
W -tryptophan
Y -tyrosine
V -valine

FIG. 3. is a schematic representation of Type I, Type II and Type III changes in *eryA* and structures of corresponding novel polyketides

10 produced. Δ69 (Type I) and Δ33 (Type II) represent in-frame deletions of the base pairs in the DNA segments corresponding to the KR of module 2 and the β-ketoacyl ACP synthase of module 2, respectively. Insertion of a complete copy of module 4 within module 1 is also shown. Production of 11-epifluoro-15-norerythromycin in strain that carries Δ33 occurs when substrate analog (2S,3S,4S,5S)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid-ethyl thioester is fed.

FIG. 4 illustrates the restriction site coordinates of cosmid pR1 5' to the sequence of *eryA* (Fig 2).

20

25

30

35

Detailed Description of the Invention

For the purposes of the present invention as disclosed and claimed herein, the following terms are defined.

The term "polyketide" as used refers to a large and diverse class of natural products, including antibiotics, pigments, and immunosuppressants. Antibiotics include, but are not limited to anthracyclines, tetracyclines, polyethers, ansamycins, macrolides of different types (polyenes and avermectins as well as classical macrolides such as erythromycins).

The term "polyketide-producing microorganism" as used herein includes any Actinomycetales which can produced a polyketide. Examples of Actinomycetes that produce polyketides include but are not limited to Micromonospora rosaria, Micromonospora megalomicea, Sacharapolyspora erythraea, Streptomyces antibioticus, Streptomyces albireticuli, Streptomyces ambofasciens, Streptomyces avermitilis, Streptomyces fradiae, Streptomyces hygroscopicus, Streptomyces tsukubaensis, Streptomyces griseus, Streptomyces mycarofasciens, Streptomyces platensis, Streptomyces venezuelae, Streptomyces

25

35

violaceoniger, and various Actinomadura, Dactylosporangium and Nocardia strains that produce polyether type of polyketides.

The term "polyketide synthase" as used herein refers to the complex of enzymatic activities responsible for the biosynthesis of polyketides which include but are not limited to β -ketoreductase, dehydratase, acyl carrier protein, enoylreductase, β -ketoacyl ACP synthase, and acyltransferase.

The term "extender" as used herein refers to a coenzyme A thioester of a dicarboxylate which is incorporated into a polyketide by a polyketide synthase.

The term "starter" as used herein refers to a coenzyme A thioester of a carboxylic acid which is used by the polyketide synthase as the first building block of the polyketide.

The term "eryA" as used herein refers to the genes involved in the formation of the polyketide moiety of erythromycin.

The term "condensation" as used herein refers to the addition of an extender unit out to the nascent polyketide chain and requires the action of β -ketoacyl ACP synthase, acyltransferase, and acyl carrier protein.

The term " β -carbonyl processing" as used herein refers to changes effecting the carbonyl group of the growing polyketide via β -ketoreductase, dehydratase, and enoylreductase.

The term "module" as used herein refers to the genetic element encoding one condensation step, as defined above, and one β -carbonyl processing step, as defined herein.

The term "Type I change" as used herein refers to changes in DNA sequence which will result in the production of polyketide rings of length identical to that of 6-deoxyerythronolide A, but with altered functional groups at specific ring positions.

The term "Type II change" as used herein refers to alterations which will result in the production of macrolide rings only when fed exogenously with substrate analogs, e.g.thioesters of appropriate acyl compounds of various length. Thus Type II mutants are erythromycin non-producing (Ery⁻) mutants. The structure of the resulting macrolides will depend on the substrate employed.

The term "Type III change" as used herein refers to alterations which will result in the biosynthesis of macrolide rings of length reduced (deletion) or increased (insertion) by two carbon units, or macrolide rings altered in specific portions of the chain (replacement).

In its broadest sense, the present invention entails a general procedure for producing novel polyketide structures in vivo by selectively altering the genetic information of the organism that naturally produces a related polyketide. A set of examples described herein are a series of novel polyketides that make use of the genetic information for the biosynthesis of the polyketide portion of the macrolide antibiotic erythromycin. The organization of the segment of the Saccharapolyspora erythraea chromosome, designated eryA, and the corresponding polypeptides which it encodes that determine the biosynthesis of the polyketide segment of erythromycin, are shown in FIG. 1. It is seen that 10 eryA is organized in modules, as shown, and that each module takes care of one condensation step, through the action of the β -ketoacyl ACP synthase specified within, wherein an extender unit, methylmalonyl CoA, is added first to the starter unit, propionyl CoA, and then to the successively growing acyl chain. The precise succession of elongation 15 steps is dictated by the genetic order of the six modules: module 1 determines the first condensation; module 2, the second; module 3, the third, and so on until the sixth condensation step has occurred. Furthermore, the processing of the growing chain after each condensation is also determined by the information within each module. Thus $\beta\text{-}$ 20 ketoreduction of the β-carbonyl takes place after each step except for step 3, as determined by the presence of a functional β -ketoreductase in all modules except module 3, whereas dehydration and enoylreduction only take place after the fourth extender unit is added to the growing acyl chain, as determined by the presence of dehydratase and enoylreductase in 25 module 4. Furthermore, the choice of the correct enantiomer (2R or 2S) of methylmalonyl-CoA as the extender unit employed at each condensation is specified by the acyltransferase function determined by each module (FIG. 1C).

In the present invention, novel polyketide molecules of desired structure are produced by the introduction of specific genetic alterations of the eryA sequence into the Sac. erythraea chromosome. The complete nucleotide sequence of the eryA segment of the Sac. erythraea chromosome and the sequence of the corresponding polypeptides are shown in FIG. 2. Three types of alterations to the eryA DNA sequence are described: (i) those inactivating a single function in a module which does not arrest acyl chain growth (β-ketoreductase, dehydratase, or enoylreductase); (ii) those inactivating a single function in a module

10

15

which does arrest chain growth (\beta-ketoacyl ACP synthase, acyltransferase or acyl carrier protein); and (iii) those affecting an entire module (deletion, insertion, or replacement). The novel polyketides produced by strains carrying these types of mutations can be classified accordingly. Type I changes will result in the production of polyketide rings of length identical to that of 6-deoxyerythronolide A, but with altered functional groups at specific ring positions. Strains carrying type II alterations will result in the production of macrolide rings only when fed exogenously with substrate analogs, e.g.thioesters of appropriate acyl compounds of various length. Thus Type II mutants are erythromycin non-producing (Ery⁻) mutants. The structure of the resulting macrolides will depend on the substrate employed. Type III changes will result in the biosynthesis of macrolide rings of length reduced (deletion) or increased (insertion) by two carbon units, or macrolide rings altered in specific portions of the chain (replacement). A schematic representation of some examples of Type I, Type II and Type III alterations in eryA and the corresponding novel polyketides produced in hosts that carry such alterations is shown in FIG. 3.

In the examples described herein, specific mutations in the eryA region of the Sac. erythraea chromosome are introduced by a simple two-20 step approach: 1) introduction of a specified change in a cloned DNA segment; 2) exchange of the wild type allele with the mutated one. Step 1 requires standard recombinant DNA manipulations employing E. coli as the host. Step 2 requires one or more plasmids out of the several E. coli-25 Sac. erythraea shuttle vectors available and a simple screening procedure for the presence of the colony carrying the altered gene. Two methods are used to introduce the altered allele into the chromosome to replace the wild type allele. The first employs gene replacement, described in Examples 7, 11, 15, 19 and 24, wherein the gene to be altered, along with adjacent upstream and downstream DNA, is mutated and cloned into a 30 Sac. erythraea non-replicating vector. The plasmid carrying the altered allele is then introduced into the host strain by transformation of protoplasts employing selection for a plasmid marker. Since the plasmid does not replicate, regenerated cells that carry the marker have undergone a single homologous recombination between one of the two segments 35 flanking the mutation on the plasmid and its homologous counterpart in the chromosome. Some of the colonies that have subsequently lost the marker will have undergone a second recombination between the other

15

plasmid borne adjacent DNA segment and its homologous chromosomal counterpart resulting in the retention of the mutation in the chromosome, replacing the normal allele with the mutant one. The second method to introduce an altered allele into the chromosome employs gene conversion, described in Examples 37 and 43. In this method, an Ery Sac. erythraea strain carrying a deletion of a specified region of the eryA segment of the chromosome is used as a host. Into a Sac. erythraea multicopy plasmid that carries a selectable marker is cloned the wild type counterpart (segment 1) of the eryA segment mutant in the host. Subsequently, the desired homologous or heterologous DNA segment to be introduced (segment 2) is cloned within the portion of segment 1 which is deleted in the mutant strain. The resulting plasmid is then introduced into the host employing selection for the marker. Among the transformants will be a population that have integrated segments 1 and 2 from the plasmid by the process of gene conversion which can be verified by examination of the DNA among the colonies that have recovered the ability to produce erythromycin.

Two examples each of Types I, II and III alterations to the eryA DNA sequence and the resultant novel polyketides produced are described in the examples described herein. Examples 1 through 8, 9 through 12 and 13 20 through 16 describe the construction and effect of three Type I mutants. Examples 17 through 22 and 23 through 27 describe the construction of two Type II mutants and the effects of feeding two different synthetic substrates to the mutant strains. Examples 28 through 38 and 39 through 44 outline the steps in constructing Type III changes and their respective 25 effects on the structure of the novel polyketides produced. In Examples 1 through 7 a plasmid that contains a substantial deletion of the segment of the gene corresponding to the b-ketoreductase of module 5 is created, the altered gene is inserted into the Sac. erythraea chromosome to replace the wild type allele and the new strain carrying the altered gene is identified 30 and isolated. In Example 8, the new strain is fermented and the novel polyketide 5-oxo-5,6-dideoxy- 3α -mycarosyl erythronolide B that results from the introduction of the mutant allele is isolated. In Examples 9 through 11, a mutation is introduced into the β -ketoreductase of module 2 and the mutated allele is then used to replace the wild type allele in the 35 chromosome. In Example 12, the strain carrying the altered allele is fermented and the novel compound 11-oxo-11-deoxyerythromycin A is isolated. Similarly, in Examples 13 through 16 a mutation is introduced

into the dehydratase of module 4 and the mutated allele is then used to replace the wild type allele in the chromosome. The strain carrying this altered allele is then fermented and the novel products 7hydroxyerythromycin A and 6-deoxy-7-hydroxyerythromycin A are isolated. In Examples 17 through 21, a mutation is made in the DNA 5 corresponding to the $\beta\text{-ketoacyl-ACP}$ synthase of module 1 and introduced into the chromosome to replace the wild type allele. This mutation has the effect of arresting the synthesis of the polyketide chain and results in the Ery phenotype. The synthetic substrate (2S,3R,4S,5S)3,5-dihydroxy-2,4dimethylhexanoic acid-ethyl ester is then made and fed to the mutant 10 resulting in the production of the novel compound (14S,15S)14(1hydroxyethyl)erythromycin. Similarly, in Examples 22 through 24, a mutation is created in the β -ketoacyl-ACP synthase of module 2 and introduced into the chromosome to replace the wild type allele. In Example 25 and 26, the synthetic substrate (2S,3S,4S,5S)2,4-dimethyl-3-15 fluoro-5-hydroxyhexanoic acid-ethyl thioester is made and fed to the module 2 β-ketoacyl-ACP synthase mutant and the resulting novel compound 11-epifluoro-15-norerythromycin is isolated. In Examples 27 through 38, a copy of the DNA sequence corresponding to module 4 is introduced into the deleted segment of the β -ketoacyl-ACP synthase of

20 introduced into the deleted segment of the β-ketoacyl-ACP synthase of module 1 resulting in the production of the novel compound 14(1-propyl)erythromycin. In Examples 40 through 44, a copy of the DNA sequence corresponding to module 5 is introduced into the deleted segment of the β-ketoacyl ACP synthase of module 1 resulting in the

production of the novel compound 14[1(1-hydroxypropyl)]erythromycin.

GENERAL METHODS

30 Materials, Plasmids and Bacterial Strains

35

Restriction endonucleases, T4 DNA ligase, nick-translation kit, competent <u>E</u>. <u>coli</u> DH5α cells , X-gal, IPTG, and plasmids pUC19 and pUC12 are purchased from Bethesda Research Laboratories (BRL), Gaithersburg, MD. [α-32P]dCTP and Hybond N are from Amersham Corp., Chicago, IL. Seakem LE agarose and Seaplaque low gelling temperature agarose are from FMC Bioproducts, Rockland, ME. <u>E</u>. <u>coli</u> K12 strains carrying the <u>E</u>. <u>coli-Sac.</u> shuttle plasmids pWHM3 or pWHM4 (Vara et al., <u>I</u>. <u>Bacteriol.</u>, 171: 5872 (1989)) or the cosmids pS1 (Tuan et al., <u>Gene</u>, 90: 21 (1990)) and

Sac. erythraea strain NRRL2338 have been deposited in the culture collection of the Agricultural Research Laboratories, Peoria, IL and are available under the accession numbers NRRL XXXX, respectively. Staphylococcus aureus ThR (thiostrepton resistant) is obtained by plating 108 cells of S. aureus on agar medium containing 10 mg/ml thiostrepton and picking a survivor after 48 hr growth at 37°C. Thiostrepton is obtained from Squibb-Bristol Myers, New Brunswick, NJ. All other chemical and reagents are from standard commercial sources unless specified otherwise.

10

35

5

DNA Manipulations

Standard conditions (Maniatis et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982) are employed for restriction endonuclease digestion, agarose gel-electrophoresis, nick translation of DNA to make ³²P-labeled probes, 15 DNA ligation, and transformation of \underline{E} . \underline{coli} employing selection for ampicillin resistance (ApR) on LB agar plates. Plasmid DNA is isolated from minipreps of \underline{E} . coli transformants by the boiling method (Maniatis et al., 1982, supra). DNA fragments are recovered from low melting agarose gels using the method of Langridge et al., 1980. Total DNA from 20 Sac. erythraea strains is prepared according to described procedures (Hopwood et al., Genetic Manipulation of Streptomyces, A Laboratory Manual, John Innes Foundation, Norwich, U.K., 1985). DNA is transferred from agarose gels onto Hybond N following the manufacturer's instructions. Hybridizations are performed in sealed bags 25 containing 10-20 ml of [1xNET (20xNET = 3 M NaCl, 0.3 M TrisHCl, 20 mM Na₂EDTA, pH 8.0), 5XDenhardt's solution (Maniatis $\underline{\text{et al.}}$, 1982, $\underline{\text{supra}}$), 0.2 mg/ml denatured calf thymus DNA, 0.2% SDS, and 0.5- $2x10^7$ cpm of the nick-translated probe] for 16-20 hr at 65 °C. Filters are washed three times in 1xNET/0.1% SDS for 20 min each at room temperature, and once in 30 0.05xNET/0.1% SDS for 20 min at 70 °C. Filters are reused as described (Donadio et al., 1990).

Amplification of DNA fragments

Synthetic deoxyoligonucleotides are synthesized on an ABI Model 380A synthesizer (Applied Biosystems, Foster City, CA) following the manufacturer's recommendations. Amplification of DNA fragments is performed by the polymerase chain reaction (PCR) employing a Coy

15

20

25

30

25

thermocycler. Reactions contain 100 pmol c ch primer, 1 µg of template DNA (cosmid pS1 carrying the *eryA* segmen from Sac. erythraea strain NRRL 2338), and 2.5 units of Thermus aquaticus DNA polymerase in a 100 ml volume of PCR buffer [50 mM KCl, 10 mM TrisHCl (pH 8.0) 2 mM MgCl₂, 0.01% gelatin) containing 200 mM of the 4 dNTPs. The above reagents are from Perkin Elmer Cetus, Norwalk, CT. The reaction mixture is overlaid with a drop of paraffin oil and subjected to 30-50 cycles. Each cycle consists of one 94 °C, one 55 °C and one 72 °C period, each of the duration of 3 min. The progress of the amplification is monitored by agarose gel-electrophoresis. The PCR primers described in the examples below are derived from the nucleotide sequence of *eryA* of FIG. 2.

Gene replacement and gene conversion

Protoplasts of Sac. erythraea strains are prepared and transformed with miniprep DNA isolated from <u>E</u>. <u>coli</u> according to published procedures (Yamamoto et al., 1986). Integrative transformants, in the case of pWHM3 derivatives, are selected after one round of non-selective growth of the primary Th^R transformants as described by Weber et. al, <u>Gene</u>, 68: 173 (1988). Loss of the Th^R phenotype is monitored by plating serial dilutions of a Th^R integrant on non-selective medium, followed by replica-plating on thiostrepton-containing medium. Th^S (thiostrepton-sensitive) colonies arise at a frequency of 10⁻² (Donadio et al., 1990). The retention of the mutant allele is established by Southern hybridization of a few Th^S colonies.

A few hundred Th^R colonies obtained by transformation of an *eryA* strain with pWHM4 derivatives are screened for antibiotic production by the agar-plug assay employing *Staphylococcus aureus* as Th^R organism as described (Tuan et a<u>l.,Gene</u>, 90: 21 (1990)). The frequency of gene conversion between a 5 kb segment of homologous sequence and a strain carrying a small deletion is >25% (Tuan et a<u>l.,Gene</u>, 90: 21 (1990)). Colonies found to produce antibiotic activity are inoculated in SGGP (Yamamoto <u>et al., 1986</u>), protoplasts are prepared, and the regenerated protoplasts are scored for loss of the plasmid by replica-plating on non-selective medium. Th^S colonies are then rechecked for antibiotic production, and six

35 producers are analyzed on Southern blots.

Fermentation

Sac. erythraea cells are inoculated into 100 ml SCM medium (1.5% soluble starch, 2.0% Soytone [Difco], 0.15% Yeast Extract [Difco], 0.01% CaCl₂) and allowed to grow at 32°C for 3 to 6 days. The entire culture is then inoculated into 10 liters of fresh SCM medium. The fermenter is operated for a period of 7 days at 32°C maintaining constant aeration and pH at 7.0. After fermentation is complete, the cells are removed by centrifugation at 4°C and the fermentation beer is kept in the cold until further use.

The present invention will now be illustrated, but is not intended to be limited, by the following examples:

Example 1 Construction of plasmid pABX9

15

10

5

The 9.6 kb <u>BamHI-XhoI</u> segment comprised between sequence coordinates 21.96 and 31.52 was isolated from cosmid pS1 and ligated to <u>SalI</u>-digested pUC19 DNA. The resulting mixture contained the desired plasmid pABX9.

20

Example 2 Construction of E. coli K12 DH5α/pABX9

Approximately 10 ng of plasmid pABX9, prepared as described in 2.5 Example 1, were transformed into <u>E</u>. <u>coli</u> K12 DH5α and a few of the resulting white Ap^R colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. One colony was found to carry pABX9, as verified by the observation of fragments of 3.93, 3.39, 2.01, 1.56, 0.87, and 0.48 kb in size upon agarose gel electrophoresis after <u>Sma</u>I digestion of the plasmid.

<u>Example 3</u> Construction of plasmid pABX9DN

Plasmid pABX9, isolated from \underline{E} . \underline{coli} K12 DH5 α /pABX9, was digested with \underline{NcoI} and then treated with T4 DNA ligase. The resulting mixture contained the desired plasmid pABX9DN.

Example 4 Construction of E. coli K12 DH5a/pABX9DN

Approximately 10 ng of plasmid pABX9DN, prepared as described in Example 3, were transformed into <u>E</u>. <u>coli</u> K12 DH5α and a few of the resulting white Ap^R colonies that appeared on the LB-agar plates containing X-gal and ampicillin were analyzed for their plasmid content. Colonies carrying pABX9DN exhibited a single <u>Nco</u>I fragment of 11.5 kb visible by agarose gel electrophoresis, confirming that the 813 bp <u>Nco</u>I - <u>Nco</u>I fragment from pABX9 has been deleted in pABX9DN.

Example 5 Construction of plasmid pABX95DN

Plasmid pABX95DN was digested with <u>Eco</u>RI and <u>Hin</u>dIII and ligated to pWHM3 digested with the same two enzymes. The resulting mixture contained the desired plasmid pABX95DN.

Example 6 Construction of E. coli K12 DH5α/pABX95DN

Approximately 10 ng of plasmid pABX95DN, prepared as described in Example 5, were transformed into E. coli K12 DH5a and a few of the resulting white Ap^R colonies that appeared on the LB-agar plates

2.5 containing X-gal and ampicillin were analyzed for their plasmid content. Colonies carrying pABX95DN exhibited fragments of 8.8 and 7.2 kb visible in agarose gels after EcoRI and HindIII digestion.

Example 7 3 0 Construction of Sac. erythraea AKR5 carrying the eryAKR5 allele by gene replacement

Approximately 1 mg of plasmid pABX95DN, isolated from <u>E</u>. <u>coli</u> K12 DH5α/pABX95DN, was transformed into *Sac. erythraea* NRRL 2338 and stable Th^R colonies were isolated. Serial dilutions of one of these colonies were screened for the loss of the antibiotic resistance marker and total DNA from 5 Th^S colonies as well as from untransformed *Sac. erythraea* NRRL 2338 was digested with <u>Ss</u>tI and analyzed by Southern

hybridization employing the 0.8 kb <u>SalI</u> fragment between sequence coordinates 24.26 and 25.06 (from pABX9) as probe. Whereas NRRL 2338 showed one <u>SstI</u> band of 3.7 kb that hybridized to the probe, samples from four of the Th^S strains exhibited a <u>SstI</u>-hybridizing band of 6.1 kb indicating the presence of the mutant allele. One of these colonies was kept and designated strain AKR5. It carries a deletion of 813 bp in the KR5 segment of *eryA* and is designated the <u>eryAKR5</u> allele.

Example 8

10 <u>Isolation, purification and properties of 5-oxo-5,6-dideoxy-3-a-mycarosyl</u> <u>erythronolide B from Sac. erythraea AKR5</u>

A 10-liter fermentation of Sac. erythrea AKR5 carrying the eryAKR5 allele in a Biolafitte fermentor using SNC Media. The fermentor was inoculated with 100 ml of a 3 day old seed. The pO2 was initially 80 ppm 15 and the temperature was maintained at 32°C. The pH was controlled to 7.0 ± 0.2 by addition of propionic acid or potassium hydroxide as needed. At harvest (3 days), the whole broth was extracted three times with 4-liter portions of ethylacetate. The combined extracts were concentrated under reduced pressure and the residue was chromatographed on a column (50 \times 20 5 cm) of silica gel packed and loaded in toluene and eluted with a stepwise gradient of increasing concentration of isopropanol in toluene. Fractions were analyzed by TLC and spots were detected by spraying with anisaldehyde sulfuric acid spray reagent and heating. A major component giving blue colored spots eluted with approximately 7% isopropanol. 25 Fractions containing this material were combined and concentrated to a residue (800 mg). This was further chromatographed on a column (100 \times 3 cm) of Sephadex LH-20 in chloroform-heptane-ethanol, 10:10:1, v/v/v. Fractions were analyzed as above, early fractions (9-13) yielded 5,6-dideoxy-3-a-mycarosyl-5-oxoerythronolide B (45 mg) which was crystallized from . 30 heptane/ethylacetate mixture to mp 163-164 °C.

CMR spectrum in CDCl3 (ppm downfield from TMS)

8.6	37.9	70.0
9.9	38.7	76.2
9.9	40.4	76.4
10.4	40.7	80.4
14.5	43.3	100.4
15.2	45.8	175.8
17.1	46.8	210.8
17.7	48.9	217.7
25.3	66.5	
25.5	69.4	

Structure was determined by single crystal X-ray diffraction.

Later fractions (15-17) yielded 5,6-dideoxy-5-oxoerythronolide B (10 mg) and still later fractions yielded 5,6-dideoxy-6,6a-epoxy-5-oxoerythronolide B (2.8 mg).

1 0 <u>Example 9</u> <u>Construction of plasmid pALeryAKR2</u>

contains the desired plasmid pALeryAKR2.

The 1.3 kb DNA segment comprised between coordinates 8.63-9.93
(fragment 1) is amplified by PCR employing two oligodeoxynucleotides, 1a
1 5 (5'-GGGAGCATGCTCTCGGTGCGCGGCGGCGGCGCGC-3') and 1b (5'GCCCTGCAGCGCGTACTCCGAGGTGGCGGT-3'). Similarly, the 1.3 kb
DNA segment between coordinates 9.99-11.26 (fragment 2) is PCRamplified employing primers 2a (5'TGGTCTGCAGGCGAGGCCGGACACCGAGG-3') and 2b (5'20 GGAAGAAGTCAAAGTTCCTCGGTCCCTTCT-3'). After digestion with
SphI + PstI (fragment 1) and PstI + EcoRI (fragment 2), the two fragments
are ligated to EcoRI + SphI-digested pWHM3. The resultant mixture

بند

4

Example 10 Construction of E. coli K12 DH5a/pALeryAKR2

Approximately 10 ng of plasmid pALeryAKR2, prepared as described in Example 9, are transformed into \underline{E} . \underline{coli} K12 DH5 α , and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryA2KR2, 9.8 kb in size, and carrying a 2.6 kb $\underline{EcoRI-SphI}$ insert with an internal \underline{PstI} site, is verified by \underline{SalI} digestion (fragments at 2.91, 2.21, 1.61, 1.42, 1.08, 0.29, 0.12 and 0.10 kb are released, visible by agarose gel electrophoresis). pALeryAKR2 contains an in-frame deletion of 102 base pairs of the corresponding segment of the wild type eryA chromosomal DNA. The cloned segment in pALeryAKR2 is designated the $\underline{eryAKR2}$ allele.

15

5

10

Example 11 Construction of Sac. erythraea AKR2 carrying the eryAKR2 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKR2, isolated from \underline{E} . \underline{coli} 20 K12 DH5 α /pALeryAKR2, is transformed into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. 25 erythraea NRRL2338 is digested with PstI and with SalI and is then examined by Southern hybridization using the 2.6 kb EcoRI-SphI insert from pALeryAKR2 as probe. Whereas NRRL2338 contains a 39 kb PstI hybridizing band, colonies in which the mutation in KR2 has been introduced (strain AKR2) exhibit two bands of approximately equal 30 intensity, one at 27 kb and the other at 12 kb. The SalI digest, with bands at 1.04, 0.75, 0.29, 0.12 and 0.10 kb common to NRRL2338 and AKR2, but with the 1.16 kb band in NRRL2338 replaced by the 1.06 kb band in AKR2, confirms that the only change introduced into strain AKR2 is the deletion of the 102 bp segment from KR2, resulting in a strain carrying the 35 eryAKR2 allele.

10

15

20

25

Example 12 Isolation and purification of 11-deoxy-11-oxoerythromycin A

The fermentation beer of strain AKR2, cooled to 4°C is adjusted to pH 8.0 and is extracted sequentially with three equal volumes of methylene chloride. The combined methylene extracts are concentrated to an oily residue and partitioned between heptane and methanol. The methanol layer is removed, washed once with heptane and concentrated to a residue. The residue is digested in methylene chloride and washed once with potassium phosphate buffer pH 7.8 and once with water. The methylene chloride layer is concentrated to a residue and digested in the lower phase (1:1:1, v/v/v) of a carbon tetrachloride; methanol; aqueous phosphate buffer (0.05 M, pH 7.0) system and chromatographed on an Ito Coil Planet Centrifuge in the same system. Fractions containing the desired 11-oxo-11-deoxyerythromycin A were combined, concentrated, digested in methylene chloride, washed well with water and concentrated on rotary evaporator under reduced pressure to yield 11-deoxy-11oxoerythromycin A as an off-white solid froth. Its identity is confirmed by comparison with antibiotic L53-18A. 11-Deoxy-11-oxoerythromycin A is dissolved in tetrahydrofuran and the solution is diluted with an equal volume of water. This is then acidified to pH 4.0 and allowed to stand at room temperature for 4 hours. The pH is adjusted to 9.0 and the solution is diluted with an equal volume of water and extracted with two volumes of methylene chloride. The combined methylene chloride extracts were evaporated to dryness under reduced pressure to yield antibiotic L53-18A as a white solid.

Example 13 Construction of plasmid pALeryADH4

30

35

Primers 3a (GCGCGAGCTCGACGACCAGGGCGGCATGGT) and 3b (GGTGGCATGCTGCGACCACTGCGCGTCGGC) are used to PCR-amplify the 1.05 kb *eryA* segment of the <u>Sac. erythraea</u> chromosome between sequence coordinates 18.47-20.07 (fragment 3), and primers 4a (AGCTGCATGCTCTGGACTGGGGACGGCTAG) and 4b

(CGCGGGATCCCACGCCGATACCG) are used to amplify the 1.35 kb segment between sequence coordinates 20.58-21.96 (fragment 4) as described in Example 1. Fragment 3 and 4, after digestion with <u>SstI + SphI</u>

and with <u>SphI + BamHI</u>, respectively, are ligated to <u>SstI -, BamHI-digested</u> pWHM3. The resulting ligation mixture contains the desired plasmid pALeryADH4.

5

20

Example 14 Construction of E. coli K12 DH5a/pALeryADH4

Approximately 10 ng of pALeryADH4, prepared as described in Example 13, are transformed transformed into <u>E</u>. <u>coli</u> K12 DH5α, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryADH4, 9.6 kb in size, is verified by <u>SphI + Eco</u>RI digestion (fragments at 7.2, 1.35 and 1.05 kb are released). pALeryADH4 carries a 498 base pair in-frame deletion of the corresponding segment of the wild type *eryA* DNA. The cloned segment in pALeryADH4 is designated the <u>eryADH4</u> allele.

Example 15 Construction of Sac. erythraea ADH4 carrying the eryADH4 allele by gene replacement

Approximately 1 mg of plasmid pALeryADH4, isolated from E. coli K12 DH5α/pALeryADH4, is used for transformation into Sac. erythraea protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and 25 $\ensuremath{\mathsf{six}}\xspace \ensuremath{\mathsf{Th}}\xspace^{\ensuremath{\mathsf{S}}}\xspace$ colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. erythraea NRRL2338 is digested with SphI and with SstI and examined by Southern hybridization using the 2.4 kb SstI-BamHI insert from pALeryADH4 as probe. Strains in which the wild type allele has been replaced by the mutated copy show two SphI bands, one at 13.5 kb and the other at 12.4 kb, whereas the wild type strain exhibits a single band at 26 kb. The SstI pattern, with the 2.9 kb band from NRRL2338 being replaced in ADH4 by a 2.5 kb band, confirms that the 487 bp deletion 3.5 created in plasmid pALeryADH4 has been transferred into the chromosome of ADH4. Strains that carry the eryADH4 allele in place of the wild type sequence are designated Sac. erythraea ADH4.

Example 16

<u>Isolation and characterization of 7-hydroxyerythromycin A and 6-deoxy-7-hydroxyerythromycin A</u>

5 The fermentation beer of strain ADH4 is cooled to 4°C and the pH is adjusted to 5.0. The mixture is extracted once with an equal volume of methylene chloride. The pH of the aqueous layer is readjusted to 9.0 and two further methylene chloride extracts are carried out. These two extracts are combined, washed with water and concentrated to a residue. This is digested in 10 ml of the upper phase of a (3:7:5, v/v/v) mixture of hexane, 10 ethylacetate, aqueous phosphate buffer (0.05 M, pH 7.5) and chromatographed on an Ito Coil Planet Centrifuge in the same system. Fractions containing the desired 7-hydroxyerythromycin were combined, concentrated, and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide solution. Fractions containing the desired 6-15 deoxy-7-hydroxyerythromycin were combined, concentrated, and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide solution. The methylene chloride layers are washed with water and then concentrated to yield the desired 7-hydroxyerythromycin A and 20 6-deoxy-7-hydroxyerythromycin A as white foams.

Example 17 Construction of plasmid pALeryAKS1

- The 1.4 kb segment of *eryA*, between sequence coordinates 1.11-2.54 (fragment 5) and the 1.5 kb segment between sequence coordinates 2.88-4.37 (fragment 6) are PCR-amplified using primers 5a (TGCAGAATTCGCTGGCCGCTCTTGGCGCT) and 5b (GAGAGCTGCAGCATGAGCCGCTGCTGCGGG), and 6a
- 3 0 (CATGCTGCAGGACTTCAGCCGGATGAACTC) and 6b (GAGGAAGCTTCCAGCCGGTCCAGTTCGTCC), respectively, as described in Example 9. After digestion with <u>EcoRI + PstI</u> (fragment 5) and <u>PstI + HindIII</u> (fragment 6), the two fragments are ligated to <u>EcoRI + HindIII-digested pWHM3</u>. The resulting mixture contains the desired plasmid
- 35 pALeryAKS1.

ŝ

20

Example 18 Construction of E. coli K12 DH5a/pALeryAKS1

Approximately 10 ng of pALeryAKS1, prepared as described in

5 Example 17, are transformed into <u>E</u>. <u>coli</u> K12 DH5α, and a few of the
resulting white ApR colonies that appear on the LB-agar plates containing
X-gal and ampicillin are analyzed for their plasmid content. The identity
of plasmid pALeryAKS1, 10.1 kb in size, is verified by digestion with <u>PstI</u>
+ <u>HindIII</u> (fragments of 8.6 and 1.5 kb are observed by agarose gel

10 electrophoresis) and with <u>SalI</u> (fragments of 2.93, 2.21, 1.42, 1.37, 0.86, 0.54,
0.27, 0.14, 0.13, and 0.10 kb are observed). pALeryAKS1 carries an in-frame
deletion of 282 base pairs of the corresponding wild type *eryA* DNA. The
cloned insert in plasmid pALeryAKS1 is designated the <u>eryAKS1</u> allele.

Example 19

Construction of Sac. erythraea AKS1 carrying the eryAKS1 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKS1, isolated from E. coli K12 DH5 α /pALeryAKS1, is used for transformation into Sac. erythraea 20 protoplasts and stable ThR colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six ThS colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six ThS colonies and from untransformed Sac. erythraea NRRL2338 is digested with PstI and with 25 SmaI and examined in Southern hybridization employing the 2.9 kb EcoRI-HindII insert from pALeryAKS1 as probe. Colonies in which the wild type allele has been replaced by the mutated copy (strain AKS1) show two PstI bands, one at 34.5 and the other at 4.4 kb, whereas the wild type strain exhibits a single band at 39 kb. The SmaI pattern, with the 2.9 kb 30 band from NRRL2338 being replaced in AKS1 by a 2.6 kb band, confirms that the 282 bp created in plasmid pALeryAKSI has been transferred into strain AKS1. Strains that carry the eryAKS1 allele are designated Sac. erythraea AKS1.

15

10

15

20

30

35

ŝ

3

Example 20

Synthesis of (2S,3R,4S,5S)3,5-dihydroxy-2,4-dimethylhexanoic acid n-butyl thioester

A convenient source of this compound in chiral purity is the antibiotic oleandomycin. Oleandomycin (5 g) is dissolved in an aprotic solvent such as toluene and treated with diazabicyclo[5.4.0]undecene-5 (1 g) and heated for one hour. The resulting solution is poured into iced water, agitated well and the organic layer is drawn off and concentrated to a residue. The residue is digested in methylene chloride and treated exhaustively with a solution of ozone. The resulting ozonide is oxidatively decomposed with dilute hydrogen peroxide in sufficient aqueous ethanol to yield a monophasic mixture. This is further diluted with water and made 0.1 N with sodium hydroxide. The mixture is warmed for one hour at 70°C and then cooled before being acidified to pH 2.5 with dilute sulfuric acid. The mixture is then exhaustively extracted with methylene chloride. The combined extracts are concentrated to an oily residue and the desired lactone is recovered by chromatography on silica gel eluted with a gradient of tolueneisopropanol.

The δ -lactone is converted to the butyl thioester before feeding to Sac. erythrea AKS1 by refluxing with n-butylthiol in the presence of a catalytic amount of triethylamine.

2.5 <u>Example 21</u> Isolation of (14S,15S)14(1-hydroxyethyl)erythromycin A

The fermentation broth of AKS1 is cooled to 4°C and adjusted to pH 4.0 and extracted once with methylene chloride. The aqueous layer is readjusted to pH 9.0 and extracted twice with methylene chloride and the combined extracts are concentrated to a solid residue. This is digested in methanol and chromatographed over a column of Sephadex LH-20 in methanol. Fractions are tested for bioactivity against a sensitive organism, such as Staphylococcus aureus ThR, and active fractions are combined. The combined fractions are concentrated and the residue is digested in 10 ml of the upper phase of a solvent system consisting of n-heptane, benzene, acetone, isopropanol, 0.05 M, pH 7.0 aqueous phosphate buffer (5:10:3:2:5, v/v/v/v/v), and chromatographed on an Ito Coil Planet

خ

Centrifuge in the same system. Active fractions are combined, concentrated and partitioned between methylene chloride and dilute ammonium hydroxide (pH 9.0). The methylene chloride layer is separated and concentrated to yield the desired product as a white foam.

5

Example 22 Construction of plasmid pALeryAKS2

Primers 7a (CGCCCGAATTCGAGGCGCTGGGCGCCCGGAC) and 7b

(CCACCTGCAGCGCGGGACCTTCCAGCCCC), and primers 8a

(GTGGGTCGCTGCAGACGGTGACTGCGG) and 8b

(GGTCAAGCTTCGTCGGCGAGCAGCTTCTC) are used to PCR-amplify

the 1.45 kb eryA segment between sequence coordinates 5.71-7.16

(fragment 7) and the 1.5 kb eryA segment between sequence coordinates

7.22-8.70 (fragment 8), respectively. After digestion with EcoRI + PstI

(fragment 7) and with <u>PstI + HindIII</u> (fragment 8), the two fragments are ligated to pWHM3 cut with <u>EcoRI + HindIII</u>. The resulting mixture contains the desired plasmid pALeryAKS2.

20

Example 23 Construction of E. coli K12 DH5a/pALeryAKS2

Approximately 10 ng of pALeryAKS2, prepared as described in Example 22, are transformed into E. coli K12 DH5α, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAKS2, 10.1 kb in size, is verified by digestion with PstI + HindIII (fragments of 8.6 and 1.5 kb are observed by agarose gel electrophoresis) and with SstII (fragments of 4.0, 2.3, 2.0, 0.72, 0.43, 0.40, 0.20, 0.18, 0.13 and 0.11 kb observed). Plasmid pALeryAKS2 carries an inframe deletion of 60 base pairs of the corresponding wild type eryA DNA. This deletion removes the active site cysteine from KS2. The cloned insert in plasmid pALeryAKS2 is designated the eryAKS2 allele.

10

15

20

AKS2.

'∄

Example 24

Construction of Sac. erythraea AKS2 carrying the eryAKS2 allele by gene replacement

Approximately 1 mg of plasmid pALeryAKS2, isolated from E. coli K12 DH5\(\alpha\)/pALeryAKS2, is used for transformation into Sac. erythraea protoplasts and stable Th^R colonies are isolated. Serial dilutions of one of these colonies are screened for loss of the antibiotic resistance marker, and six Th^S colonies are analyzed for their genotype by Southern hybridization. Total DNA from the six Th^S colonies and from untransformed Sac. erythraea NRRL2338 is digested with PstI and with SstII and examined in Southern hybridization employing the 2.9 kb EcoRI-HindII insert from pALeryAKS2 as probe. Colonies in which the wild type allele has been replaced by the mutated copy (strain AKS2) show two PstI bands, one at 34.5 and the other at 4.4 kb, whereas the wild type strain exhibits a single band at 39 kb. The SstII pattern, with the 0.78 kb band from NRRL2338 being replaced in AKS2 by a 0.72 kb band, confirms that the 60 bp created in plasmid pALeryAKS2 has been transferred into strain AKS2. Strains that carry the eryAKS2 allele are designated Sac. erythraea

Example 25

Synthesis of (2R,3R,4S,5R)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid n-butyl thioester

25

30

35

(2R,3S,4S,5R)3,5-Dihydroxy-2,4-dimethylhexanoic acid-\delta-lactone (1 g) from Example 20 is digested in 10 ml of pyridine and treated with p-toluenesulfonyl chloride (1.3 g) and allowed to stand at room temperature overnight. The mixture is poured into iced water and extracted with methylene chloride and the methylene chloride is concentrated to the crude sulfonate ester. This is digested in acetonitrile (100 ml) and heated under reflux after the addition of tetrabutylammonium fluoride (1.75 g). After 6 hours the mixture is cooled, poured over iced water (300 ml) and extracted three times with 200 ml portions of methylene chloride. The combined methylene chloride extracts were concentrated and the residue was chromatographed on a column of silica gel eluted with a stepwise gradient of isopropanol (0 to 50%) in toluene. Fractions containing (2R,3R,4S,5R)2,4-dimethyl-3-fluoro-5-hydroxyhexanoic acid_d-lactone were

combined and concentrated to a white solid. The lactone is then converted to the n-butyl thiolester by refluxing in n-butyl thiol with a catalytic amount of triethylamine. Solvent is removed and the residue is digested in DMSO before feeding to fermentations of <u>Sac. erythraea</u> AKS2.

5

30

Example 26 Isolation and purification of 11-epifluoro-15-norerythromycin A

The fermentation broth of strain AKS2 is cooled to 4°C and adjusted to pH 4.0 and extracted once with ethylacetate. The aqueous layer is 10 adjusted to pH 9.0 and extracted twice with methylene chloride and the combined extracts are concentrated to a white solid. This is chromatographed over a column of Sephadex LH-20 in a mixture of heptane, chloroform, ethanol (10:10:1, v/v/v) and fractions containing the desired product are combined and concentrated to a solid residue. This is 15 further purified by countercurrent chromatography on an Ito Coil Planet Centrifuge on a system composed of carbon tetrachloride; methanol; 0.05 M; pH 7.0 aqueous potassium phosphate buffer (I:1:1, v/v/v). Fractions containing the desired 11-epifluoro-15-norerythromycin were combined, and concentrated to a residue. This was digested in methylene chloride 20 and dilute (pH 9.5) ammonium hydroxide and the methylene chloride layer was separated, washed with water and concentrated to yield the desired 11-epifluoro-15-norerythromycin A as white solid.

2.5 <u>Example 27</u> Construction of plasmid pALeryAM4.1

Primers 9a (GCGCCGAATTCTCGAGACGCGTGGGAGGCA) and 9b (TTGCGGTACCAGTAGGAGGCGTCCATCGCG) are employed to PCR-amplify the 2.0 kb *eryA* segment between sequence coordinates 17.35-19.38 (fragment 9). After digestion with <u>EcoRI + KpnI</u>, fragment 9 is ligated to pUC19 cut with the same two enzymes The resulting mixture contains the desired plasmid pALeryAM4.1.

* .

Example 28 Construction of E. coli K12 DH5a/pALeryAM4.1

Approximately 10 ng of pALeryAM4.1, prepared as described in

5 Example 27, are transformed into E. coli K12 DH5a, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 4.7 kb in size, is verified by digestion with SalI (fragments of 2.8, 0.85, 0.53, 0.27 and 0.22 kb are observed by agarose gel electrophoresis).

Example 29 Construction of plasmid pALeryAM4.2

15 Primers 10a (GCTGGGATCCCGCGGGGGGGGGGGGGACAC) and 10b (CGGAACTCGGTGAGCATGCCGGGACTGCTC) are used to PCR-amplify the 2.1 kb *eryA* segment between sequence coordinates 21.94-24.00 (fragment 10). The 2.6 kb fragment KpnI(96)-BamHI(102) from cosmid clone pR1, and fragment 10 cut with BamHI + SphI, are ligated to pALeryAM4.1 cut with KpnI + SphI. The resulting mixture contains the desired plasmid pALeryAM4.2.

Example 30 Construction of E. coli K12 DH5a/pALeryAM4.2

25

30

Approximately 10 ng of pALeryAM4.2, prepared as described in Example 29, are transformed into E. coli K12 DH5a, and a few of the resulting white ApR colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.2, 9.3 kb in size, is verified by digestion with XhoI + SphI (to ensure that the entire 6.65 kb insert is released) and with SalI, with fragments of 2.8, 1.82, 1.09, 0.94, 0.85, 0.75, 0.45, 0.27, 0.22 and 0.13 kb are observed by agarose gel electrophoresis).

٠

Example 31 Construction of plasmid pALeryAM1

The 2.9 kb <u>SmaI(4)-SmaI(20)</u> fragment from cosmid clone pR1 is ligated to pUC12 cut with <u>SmaI</u>. The resulting mixture contains plasmid pALeryAM1.

Example 32 Construction of E. coli K12 DH5αa/pALeryAM1

10

15

20

Approximately 10 ng of pALeryAM1, prepared as described in Example 31, are transformed into \underline{E} . \underline{coli} K12 DH5 α , and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM1, 5.6 kb in size, is verified by digestion with \underline{SmaI} (the 2.9 kb insert is realeased) and with \underline{SphI} , with release of one 4.4 and one 1.07 kb bands. Both orientations of the insert in plasmid pALeryAM1 are useful.

Example 33 Construction of plasmid pALeryAM4.3

Plasmid pALeryAM1 is cut with <u>Xho</u>I to completion, partially with <u>SphI</u>, and the resulting 5.25 kb band, isolated from an agarose gel, is ligated to the 6.65 kb insert released from pALeryAM4.2 by <u>Xho</u>I + <u>SphI</u> digestion The resulting mixture contains the desired plasmid pALeryAM4.3.

Example 34 Construction of E. coli K12 DH5a/pALeryAM4.3

30

35

Approximately 10 ng of pALeryAM4.3, prepared as described in Example 33, are transformed into \underline{E} . \underline{coli} K12 DH5 α_z and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 11.9 kb in size, is verified by $\underline{XhoI} + \underline{SphI}$ digestion (fragments of 6.65 and 5.25 kb are visible by agarose gelelectrophoresis). Plasmid pALeryAM4.3 carries the entire eryA module 4

10

٠

inserted into the KS region of module 1. The cloned insert in pALeryAM4.3 is degnated the <u>eryAM412</u> allele.

Example 35 Construction of plasmid pALeryAM4.4

Plasmid pALeryAM4.3 is cut with <u>Eco</u>RI + <u>Hind</u>III, and the resulting 9.2 kb band, recovered from an agarose gel, is ligated to pWHM4 cut with the same two enzymes. The resulting mixture contains the desired plasmid pALeryAM4.4.

Example 36 Construction of E. coli K12 DH5\(\alpha\)/pALeryAM4.4

Approximately 10 ng of pALeryAM4.4, prepared as described in Example 35, are transformed into <u>E</u>. <u>coli</u> K12 DH5α, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM4.1, 16.5 kb in size, is verified by <u>Eco</u>RI + <u>Hind</u>III digestion, with fragments of 9.2 and 7.3 kb released. Plasmid pALeryAM4.4 carries the <u>eryAM412</u> allele on the <u>Sac</u>. <u>erythraea</u> multicopy vector pWHM4.

Example 37

25 <u>Construction of Sac. erythraea AM412 carrying the eryAM412 allele by gene conversion</u>

Approximately 1 mg of plasmid pALeryAM4.4, isolated from E. coli K12 DH5\(alpha\)/pALeryAM4.4, is used for transformation into Sac. erythraea strain AKS1 protoplasts. A few hundred transformants are screened for antibiotic production by the agar-plug assay, and one of the colonies found to produce antimicrobial activity is cured of pALeryAM4.4 by protoplast formation and regeneration as described in General Methods. Total DNA from six antibiotic-producing, ThS colonies (strain AM412)and from strain AKS1 is digested with SphI and with XhoI and the resulting Southern blot is hybridized first to the 2.9 kb insert from pALeryAM1, and then to the 2.9 kb SstI(95)-SstI(101) fragment from plasmid pALeryAM4.2. With the first probe, the SphI band at 0.8 kb in strain AKS1 is seen to be replaced by a 7.5

*

kb band in strain AM412, whereas the other two bands at 2.4 kb and 5.2 kb are unaffected. In the XhoI digest, the AKS1 band at 2.9 kb is replaced by a 9.6 kb band in AM412, with the other band at 5.2 kb conserved in both strains. Using the SstI(95)-SstI(101) fragment as probe, strain AKS1 exhibits one band at 25.5 kb and one at 17.9 kb in the SphI and XhoI digests, respectively, whereas, in addition to these bands, strain AM412 shows one SphI band at 7.5 kb and one XhoI band at 9.6 kb. In this way, it is established that the eryAKS1 allele has been converted into the eryAM412 allele in strain AM412.

10

5

Example 38 Isolation and purification of 14-(1-propyl)erythromycin A

At harvest the fermentation is adjusted to pH 9.5 and extracted twice with equal volumes of methylene chloride. The combined extracts 15 are washed once with water and concentrated to an oily residue. This is partitioned in a heptane methanol water (5:5:1, v/v/v) system and the lower layer is washed once with heptane and then concentrated to a semisolid residue. This is digested in methanol and chromatographed over a column of Sephadex LH-20 in methanol. Fractions are tested for 20 bioactivity in an agar diffusion assay on plates seeded with the macrolidesensitive strain Staphylococcus aureus ThR. Active fractions are combined and further purified by chromatography over silica gel a chloroform:methanol gradient containing 0.1% triethylamine. Fractions containing the desired 14-(1-propyl)erythromycin A are combined and 25 concentrated to yield the product as a white solid.

Example 39 Construction of plasmid pALeryAM5.1

30

35

The 4.7 kb *eryA* fragment between sequence coordinates 23.65-28.36 (fragment 11) is PCR-amplified employing primers 11a (ATGCTCGAGATCTCGTGGGAGCGCGCGA) and 11b (AGAACTCGGTGAGCATGCCCGGGCCCCA). Fragment 11, after digestion with XhoI + SphI, is ligated to the 5.25 kb fragment resulting from complete XhoI and partial SphI digestion of pALeryAM1, as in Example 33. The resulting mixture contains the desired plasmid pALeryAM5.1.

Example 40 Construction of E. coli K12 DH5α/pALeryAM5.1

Approximately 10 ng of pALeryAM5.1, prepared as described in Example 39, are transformed into \underline{E} . \underline{coli} K12 DH5 α_z and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM5.1, 9.95 kb in size, is verified by $\underline{SphI} + \underline{XhoI}$ digestion, with fragments of 5.25 and 4.7 kb released, and by \underline{SmaI} digestion where fragments of 3.39, 2.68 and 1.94 (doublet) kb are observed. Plasmid pALeryAM5.1 carries the entire eryA module 5 inserted into the β -ketoacyl ACP synthase region of module1. The cloned insert in plasmid pALeryAM5.1 is designated the $\underline{eryA512}$ allele.

15

25

5

10

ŝ

Example 41 Construction of plasmid pALeryAM5.2

Plasmid pALeryAM5.1 is cut with <u>EcoRI + HindIII</u> and the resulting 6.3 kb fragment, recovered from an agarose gel, is ligated to pWHM4 cut with the same two enzymes. The resulting mixture contains the desired plasmid pALeryAM5.2.

Example 42 Construction of E. coli K12 DH5α/pALeryAM5.2

Approximately 10 ng of pALeryAM5.2, prepared as described in Example 41, are transformed into <u>E</u>. <u>coli</u> K12 DH5α, and a few of the resulting white Ap^R colonies that appear on the LB-agar plates containing X-gal and ampicillin are analyzed for their plasmid content. The identity of plasmid pALeryAM5.2, 13.6 kb in size, is verified by digestion with <u>Eco</u>RI + <u>Hind</u>III, with fragments of 7.3 and 6.3 kb released. Plasmid pALeryAM5.2 contains the <u>eryAM512</u> allele on the <u>Sac. erythraea</u> multicopy vector pWHM4.

خ

5

Example 43 Construction of Sac. erythraea AM512 carrying the eryAM512 allele by gene conversion

Approximately 1 mg of plasmid pALeryAM5.2, isolated from E. coli K12 DH5 α /pALeryAM5.2, is used for transformation into Sac. erythraea strain AKS1 protoplasts. A few hundred transformants are screened for antibiotic production by the agar-plug assay, and one of the colonies found to produce antimicrobial activity is cured of pALeryAM5.2 by protoplast formation and regeneration as described in General Methods. Total DNA 10 from six antibiotic-producing, ThS colonies (strain AM512)and from strain AKS1 is digested with SphI and with XhoI and the resulting Southern blot is hybridized first to the 2.9 kb insert from pALeryAM1, and then to the 0.8 kb NcoI(119)-NcoI(123) fragment from plasmid pALeryAM5.1. With the first probe, the SphI band at 0.8 kb in strain AKS1 is replaced by a 5.5 kb 15 band in strain AM512, whereas the other two bands at 2.4 kb and 5.2 kb are unaffected. In the XhoI digest, the AKS1 band at 2.9 kb is replaced by a 7.6 kb band in AM512, with the other band at 5.2 kb conserved in both strains. Using the NcoI(119)-NcoI(123) fragment as probe, strain AKS1 exhibits one band at 25.5 kb and one at 17.9 kb in the SphI and XhoI digests, 20 respectively, whereas, in addition to these bands, strain AM512 shows one SphI band at 5.5 kb and one XhoI band at 7.6 kb. In this way, it is established that the eryAKS1 allele has been converted into the eryAM512 allele in strain AM512.

25

Example 44 Isolation and purification of 14[1(1-hydroxypropyl)]erythromycin A

At harvest the pH of the fermentation of AM512 is adjusted to 9.5 and the mixture is extracted twice with equal volumes of ethylacetate. 30 The combined ethylacetate extracts are washed with water, dried and partitioned in a heptane, methanol, water (5:5:1, v/v/v) system. The lower (methanolic phase) is washed with an equal volume of heptane and is concentrated to a residue. This is chromatographed on a Sephadex LH-20 column in methanol and fractions containing the desired 14[1(1-35 hydroxypropyl)]erythromycin A are concentrated and further purified by chromatography on an Ito Coil Planet Centrifuge in a system consisting of n-heptane, benzene, acetone, isopropanol, 0.65 M, pH 7.0 aqueous

÷

potassium phosphate buffer (5:10:2:3:5, v/v/v/v). Fractions containing the desired product are concentrated to a solid residue and partitioned between methylene chloride and dilute (pH 9.5) ammonium hydroxide. The organic layer is washed with water and concentrated to yield 14[1(1-hydroxypropyl)]erythromycin A as a white solid.

Although the present invention is described in the Examples listed above in terms of preferred embodiments, they are not to be regarded as limiting the scope of the invention. The above descriptions serve to illustrate the principles and methodologies involved in creating the three 10 types of mutations that can be introduced into the eryA segment of the Sac. erythraea chromosome that result in the synthesis of novel polyketide products. Although single Type I alterations, leading to the production of 5-oxo-5,6,-dideoxy-3α-mycorosyl erythronolide B, 11-oxo-11deoxyerythromycin A, 7-hydroxyerythromycin A, 7-oxo-15 7deoxyerythromycin A, 5-desosaminyl-3-oxo-3-deoxyerythronolide A, and Δ-6,7-anhydro-6-deoxyerythromycin A are specified herein, it is obvious that other Type I changes can be introduced into the eryA segment leading to novel polyketide structures. Among the additional Type I alterations that can be obtained are those in which two or more modules 20 are affected leading to the synthesis of novel polyketides. Examples of combinations of two Type I alterations leading to useful compounds include but are not limited to: mutants of the the β -ketoreductase of module 2 (KR2) and the β -ketoreductase of module 4 (KR4) leading to the 25 formation of 7,11-dioxo-7,11-dideoxyerythromycin A; mutants of KR2 and the β-ketoreductase of module 6 (KR6) leading to the formation of 3,11dioxo-3,11-dideoxy-5-desosaminylerythronolide A; mutants of KR2 and the dehydratase of module 4 (DH4) leading to the synthesis of 7-hydroxy-11-oxo-11-deoxyerythromycin A; mutants of KR2 and the enoylreductase 30 of module 4 (ER4) leading to the synthesis of Δ -6,7-anhydro-11-oxo-11deoxyerythromycin A; mutants of KR4 and KR6 leading to the synthesis of 3,7-dioxo-3,7-dideoxy-5-desosaminylerythronolide A; mutants of KR6 and DH4 leading to the synthesis of 3-oxo-3-deoxy-5-desosaminyl-7hydroxyerythronolide A; mutants of KR6 and ER4 leading to the synthesis of 3-oxo-3-deoxy-5-desosaminyl- Δ -6,7-anhydroerythronolide A. Examples 35 of combinations of three Type I alterations leading to the synthesis of novel polyketides include but are not limited to: mutants of KR2, KR4 and KR6 leading to the synthesis of 3,7,11-trioxo-3,7,11-trideoxy-5-

Š

25

30

35

desosaminylerythronolide A; mutants of KR2, KR6 and DH4 leading to the synthesis of 3,11-dioxo-3,11-dideoxy-5-desosaminyl-7-hydroxyerythronolide A; mutants of KR2, KR6 and ER4 leading to the synthesis of 3,11-dioxo-3,11-dideoxy-5-desosaminyl-D-6,7-anhydroerythronolide A. All combinations of two or three Type I mutants, the <u>Sac. erythraea</u> strains that carry said combinations and the

mutants, the <u>Sac. erythraea</u> strains that carry said combinations and the corresponding polyketides produced from said strains, therefore, are included within the scope of the present invention.

Although the Type II mutants specified herein have been constructed in the $\beta-ketoacyl$ ACP synthase of module 1 (KS1) and the $\beta-$ 10 ketoacyl ACP synthase of module 2 (KS2), other Type II mutants can be constructed in other domains to result in the synthesis of novel polyketide structures upon feeding with appropriate substrate analogs. Other Type II mutants include but are not limited to: inactivation of the either of the acyltransferases or acyl carrier proteins of module 1, or the 15 acyltransferase or acyl carrier protein of module 2, the β -ketoacyl ACP synthase, acyltransferase or acyl carrier protein of module 3, module 4 or module 5. Furthermore, compounds other than (2S,3R,4S,5S)3,5dihydroxy-2,4-dimethylhexanoic acid-ethyl thioester and (2S,3S,4S,5S)2,4dimethyl-3-fluoro-5-hydroxyhexanoic acid-ethyl thioester specified herein 20 can be synthesized and fed to strains AKS1 or AKS2 specified herein or other strains that carry other Type II mutations to result in the creation of novel polyketides that are within the scope of the present invention.

Although two examples of Type III alterations are specified herein, it is apparent to those skilled in the art that many other examples of Type III changes are possible. Strains of *Sac. erythraea* carrying changes of this type offer the very high potential for the production of novel polyketides of specified structure, since they do not require synthetic substrates as do Type II mutants and they are not limited to the formation of derivatives of erythromycin, as in the case of Type I mutants. In the embodiments of Type III mutants specified herein, we have illustrated how a second copy of a complete module can be introduced at a desired position by gene conversion to result in the synthesis of 14-(1-propyl)erythromycin A or 14-[1(1-hydroxypropyl])erythromycin A. These alterations make use of the high conservation and simultaneous lack of specificity of the β -ketoacyl ACP synthases of modules 1 and 2, thereby making possible the construction of hybrid β -ketoacyl ACP synthase functions consisting of portions of proteins derived from different modules. Those skilled in the

ŝ

10

15

30

35

art understand, therefore, that it is possible, as exemplified for KS1 and KS2, to delete a small portion of the β -ketoacyl ACP synthase of other modules and to construct strains carrying such alterations which can then be employed as hosts for introducing at the deleted β -ketoacyl ACP synthase location a second copy of any homologous module. Furthermore, as exemplified herein, it is also possible to delete any segment of eryA by ligation of two non-contiguous PCR-generated fragments and to subsequently construct strains, therefore, devoid of any or all portions of any module. Such strains deleted of a full module can be employed for reintroduction of either the same or a different module at a different location. It is possible, therefore, to determine the novel structures desired and then create a series of Sac. erythraea strains containing the corresponding arrangements of eryA modules that would produce said novel structures that are included within the scope of the present invention. Additional examples of novel compounds produced from the construction of Type III alterations include but are not limited to 11-deoxyerythromycin, resulting from the insertion of the eryA segment encoding DH4 and ER4 in module 2.

Moreover, it will also be apparent that two or more modules can be excised and introduced into various sites of the Sac. erythraea chromosome to produce novel polyketides of predicted structure such as the introduction of the eryA segment encoding DH4 and ER4 in both module 1 and module 2 to result in the production of 14(R)[1-hydroxypropyl]11-deoxyerythromycin A. All combinations, therefore, of Type III alterations and the strains of Sac. erythraea that carry said alterations as well as the polyketides produced from said strains are included within the scope of the present invention.

In addition, it is also possible to create combinations of Type I, Type II and Type III alterations and insert such alterations into Sac. erythraea to produce novel polyketides. Examples of such combinations include but are not limited to the following. The combination of a Type I alteration, such as an alteration in DH4 and a Type II alteration, such as a mutation in the KS1 to result in the formation of (14S,15S)14-[1-hydroxyethyl]-7-hydroxyerythromycin A when the strain of Sac. erythraea carrying such alterations is fed with the compound (2S,3R,4S,5S)3,5-dihydroxy-2,4-dimethylhexanoic acid ethyl ester. The combination of a Type I alteration, such as an alteration in DH4 and a Type III alteration, such as found in Sac. erythraea strain AM412, wherein a copy of the DNA segment of

ε

10

15

20

2:5

30

35

module 4 is introduced in module 1, such that the *Sac. erythraea* strain so constructed produces the compound 7-hydroxy-14-propylerythromycin A. All combinations of two or more alterations of Type I, Type II and Type III alterations, the *Sac. erythraea* strains that carry such alterations, and the polyketides produced from such strains are included within the scope of the present invention. It will also occur to those skilled in the art that novel structures can be produced by altering the specificity of the acyltransferase functions in any module. Examples include: replacement of the acyltransferase domains of modules 1, 2, 3, 4, 5, or 6 in *eryA* with those of modules 4, 4, 2, 2, 2, and 4, respectively, to result in the production of 12-epierythromycin A,10-epierythromycin A, 8-epierythromycin A, 6-epierythromycin A, 4-epierythromycin A and 2-epierythromycin A, respectively, that are included within the scope of the present invention.

It should be emphasized that the introduction of an entire eryAmodule at a different location, as exemplified for the construction of Sac. erythraea strains AM412 and AM512 in Examples 29 and 35, respectively, does not rely on homologous recombination between the incoming eryA module and the host chromosome. Rather, gene conversion of the host allele with the eryA allele residing on the multicopy plasmid requires DNA sequences homologous to the host allele flanking the incoming module. Thus, any module carrying the desired specificities, either from homologous or heterologous sources, can be employed in gene conversion of the host allele, provided that is flanked by segments of homology. It will occur to those skilled in the art, therefore, that, given the large number of natural polyketide molecules existing, a wide variety of additional novel molecules of predicted structure can be produced in Type III mutants containing an additional module of desired specificities or where an endogenous module is replaced by an exogenous one. The length of the acyl chain can be easily controlled by suitably changing the number of modules involved in its synthesis. Similarly, the introduction of keto, hydroxy, enoyl, or methylene groups at specific points along the acyl chain can be easily achieved by introducing the proper b-carbon processing functions (β -ketoreductase, dehydratase and enoylreductase) in the required modules. Exogenous modules constitute the source of specificities for starter and extender units other than those employed by Sac. erythraea for erythromycin biosynthesis, making it thereby possible to employ, for example, malonylCoA or (2R)- or (2S)ethylmalonylCoA, etc.

10

20

25

30

35

as extender units, and acetyl CoA, butyryl CoA, etc. as the starter unit. The result will be the formation of erythromycin analogs containing the desired functional groups and side chains with the desired stereochemistry. As an extension of the examples reported with eryA, the construction of a Sac. erythraea strain carrying a heterologous module inserted into eryA requires: (i) cloning of the genes from any other Actinomyces producing a polyketide with desired structural features; (ii) mapping of the modular organization of the cloned genes by low stringency hybridization and restriction analysis; (iii) locating the module carrying the desired specificities by partial sequencing; (iv) precise excision of the desired genetic element and cloning into a vector suitable for gene conversion; (v) construction and transformation of a Sac. erythraea strain suitable for gene conversion and screening for the novel compound. Any module, or portion thereof, can thus be precisely excised from the genome of a polyketide-producing microorganism and introduced into suitable Sac. erythraea strains to create a novel polyketide of predicted structure. Thus, replacement of the acyltransferase segments of modules 1, 2, 3, 4, 5,or 6 in eryA with the acyltransferase segment specific for malonyl CoA, such as can be found in the polyketide synthase genes for the synthesis of pikromycin in Streptomyces venezuelae, to result in the synthesis of 12norerythromycin A, 10-norerythromycin A, 8-norerythromycin A, 6norerythromycin A, 4-norerythromycin A and 2-norerythromycin A, respectively, that are included within the scope of the present invention. In addition, replacement of the acyltransferase segments of modules 1, 2, 3, 4, 5, or 6 in ery A with an acyltransferase specific for (2R)-ethylmalonyl CoA, such as can be found in the polyketide synthase genes for the synthesis of spiramycin in Streptomyces ambofasciens, will result in the formation of 12-homoerythromycin A, 10-homoerythromycin A, 8epihomoerythromycin A, 6-epihomoerythromycin A, 4epihomoerythromycin A and 2-homoerythromycin A, respectively, all of which are included within the scope of the present invention. Similarly, introduction of acyltransferase segments carrying desired specificities for the starter or extender unit into eryA DNA that results in the synthesis of novel compounds are included within the scope of the present invention. The erythromycin analogs produced by the method of this invention are structurally similar to known antibacterial and prokinetic agents.

It will also occur to those skilled in the art that genetic manipulations described herein need not be limited to Sac. erythraea.

ş

Ŧ

÷.

5

10

15

20

25

30

35

Suitable hosts are any other polyketide-producing Actinomyces where DNA can be precisely inserted into the chromosome. Hence, the choice of a convenient host is based solely on the relatedness of the novel polyketide to a natural counterpart so as to minimize the number of module rearrangements required for its biosynthesis. Therefore, Type I, Type II and Type III alterations can be constructed in other Actinomyces employing either endogenous or exogenous modules to produce novel polyketides employing strategies analogous to those described herein for Sac. erythraea. Thus all Type I, Type II or Type III mutations or various combinations thereof constructed in other actinomycetes according to the principles described herein, and the respective polyketides produced from such strains, are included within the scope of the present invention. Examples of polyketides that can be altered by creating Type I, Type II or Type III changes in the producing microorganisms include, but are not limited to macrolide antibiotics such as erythromycin, tylosin, spiramycin, etc.; ansamacrolides such as rifamycins, maytansines, etc.; polyketide antibiotics such as tetracycline; polyethers such as monesin, salinomycin, etc.; polyenes such as candicidin, amphothericins; immunosuppressants such as FK506, ascomycin, rapamycin, etc. and other complex polyketides such as avermectin.

Whereas the novel derivatives or modifications of erythromycin described herein have been specified as the A derivatives, such as 7hydroxyerythromycin A, 11-oxo-11-deoxyerythromycin A, 14[1(1hydroxypropyl)]erythromycin A, etc., those skilled in the art understand that the wild type strain of Sac. erythraea produces a family of erythromycin compounds, including erythromycin A, erythromycin B, erythromycin C and erythromycin D. Thus, modified strains of Sac. erythraea, such as strain AKR2, for example, would be expected to produce the corresponding members of the 11-oxo-11-deoxyerythromycin family, including 11-oxo-11-deoxyerythromycin A, 11-oxo-11-deoxyerythromycin B, 11-oxo-11-deoxyerythromycin C, and 11-oxo-11-deoxyerythromycin D. Similarly, strain AM412 would be expected to produce not only 14(1propyl)erythromycin A but also the other members of the 14(1propyl)erythromycin family including 14(1-propyl)erythromycin B, 14(1propyl)erythromycin C and 14(1-propyl)erythromycin D. Similarly, all other modified strains of Sac. erythraea described herein that produce novel erythromycin derivatives would be expected to produce the A, B, C, and D forms of said derivatives. Therefore, all members of the family of

5

10

15

20

25

30

35

Ŀ

۶,

each of the novel polyketides described herein are included within the scope of the present invention.

Variations and modifications of the methods for obtaining the desired plasmids, hosts for cloning and choices of vectors and segments of eryA DNA to clone and modify, other than those described herein that result in substantially the same strains and same products as those described herein will occur to those skilled in the art. For example, although we have described the use of the plasmids pWH3 and pWHM4 as E. coli-Sac. erythraea shuttle vectors, other vectors can be employed wherein all or part of pWHM3 or pWHM4 is replaced by other DNA segments that function in a similar manner, such as replacing the pUC19 component of pWHM3 and pWHM4 with pBR322, available from BRL, employing different segments of the pIJ101 or pJV1 replicons in pWHM3 and pWHM4, respectively, or employing selectable markers other than thiostrepton- and ampicillin-resistance. These are just few of a long list of possible examples all of which are included within the scope of the present invention. Similarly, the segments of the eryA locus subcloned into pWHM3 for generating strains AKS1, AKS2, etc. specified herein can readily be substituted for other segments of different length encoding the same functions, either produced by PCR-amplification of genomic DNA or of an isolated clone, or by isolating suitable restriction fragments from Sac. erythraea. In the same way, it is possible to create eryA strains carrying mutations functionally equivalent to those described herein by deleting different portions of the corresponding genes, by creating insertions into them, or by site-directed mutagenesis of specific nucleotide residues. Moreover, Sac. erythraea strains with mutant alleles other than the β ketoacyl ACP synthase portions of eryA can be employed as hosts for gene conversion; Type III mutants can be constructed by double reciprocal crossover as exemplified for Type I and Type II mutants rather than by the gene conversion method described herein. Additional modifications include changes in the restriction sites used for cloning or in the general methodologies described above. All such changes are included in the scope of the invention. It will also occur to those skilled in the art that different methods are available to ferment Sac. erythraea, to extract the novel polyketides specified herein, and to synthesize substrate analogs, and that all such methods are also included within the scope of the present invention.

It will be apparent that many modifications and variations of the invention as set forth herein are possible without departing from the spirit and scope thereof, and that, accordingly, such limitations are imposed only as indicated by the appended claims.

5

.

ð

Ç

5

25

30

Į,

What is claimed is:

- 1. A method for directing the biosynthesis of specific polyketide analogs by genetic manipulation of a polyketide-producing microorganism, said method comprising the steps of:
- isolating a polyketide biosynthetic gene-containing DNA sequence;
- (2) identifying enzymatic activities associated within said genecontaining DNA sequence;
- 10 (3) introducing one or more specified changes into said genecontaining DNA sequence which codes for one of said enzymatic activities resulting in an altered DNA sequence;
 - (4) introducing said altered DNA sequence into a polyketideproducing microorganism to replace the original sequence;
- 15 (5) growing a culture of the altered microorganism under conditions suitable for the formation of the specific polyketide analog; and (6) isolating said specific polyketide analog from the culture.
- 2. The method of claim 1 wherein said polyketide biosynthetic gene-20 containing DNA sequence comprises genes which encode the enzymatic activities comprising a polyketide synthase.
 - 3. The method of claim 2, wherein said polyketide synthase enzymatic activities comprise β -ketoreductase, dehydratase, acyl carrier protein, enoylreductase, β -ketoacyl ACP synthase, and acyltransferase.
 - 4. The method of claim 1 wherein said alteration which occurs in the DNA sequence results in the inactivation of one or more enzymatic activities involved in the processing of the β -carbonyl of said polyketide.
 - 5. The method of claim 4, wherein said inactivated enzymatic activities affecting the processing of the β -carbonyl of said polyketide comprise β -ketoreductase, dehydratase, and enoylreductase.
- 3 5 6. The method of claim 4 wherein said alteration in the DNA sequence results in the addition of one or more enzymatic activities involved in the β-carbonyl processing of said polyketide.

PCT/US92/00427

8

Š

- 7. The method of claim 6 wherein said additional enzymatic activities are selected from the group consisting of β -ketoreductase, β -ketoreductase and dehydratase, and β -ketoreductase, dehydratase and enoylreductase.
- 5 8. The method of claim 1 wherein said alteration occurring in the DNA segment results in the inactivation of one or more enzymatic activities involved in the condensation of carbon units to the nascent polyketide structure.
- 10 9. The method of claim 8 wherein said enzymatic activities affecting the condensation of carbon units to the nascent polyketide structure comprise β-ketoacyl ACP synthase, acyl carrier protein and acyltransferase.
- 10. The method of claim 1 wherein said alteration in the DNA15 sequence results in the change of the length of the polyketide synthesized.
 - 11. The method of claim 10 wherein said alteration results in the increase of the length of the polyketide.
- 20 12. The method of claim 11 wherein said alteration comprises the addition of DNA sequences encoding the enzymatic activities consisting of acyltransferase, acyl carrier protein and β -ketoacyl ACP synthase.
- 13. The method of claim 10 wherein said alteration results in the25 decrease of the length of the polyketide.

30

3.5

- 14. The method of claim 13 wherein said alteration consists of the deletion of a DNA segment between two sequences encoding corresponding enzymatic activities.
- 15. The method of claim 14 wherein said corresponding enzymatic activities are selected from the group consisting of β -ketoreductases, dehydratases, acyl carrier proteins, enoylreductases, β -ketoacyl ACP synthases, and acyltransferases.
- 16. The method of claim 1 wherein said alteration consists in the replacement of the DNA segment encoding an acyltransferase with a DNA segment encoding an acyltransferase of different specificity.

*

15

- 17. The method of claim 1 wherein said DNA sequence is isolated from a species from the *Actinomycetales* family.
- 5 18. The method of claim 17 wherein said DNA sequence is isolated from a genus selected from the group consisting of Actinomyces, Dactylosporangium, Micromonospora, Nocardia, Sac., Streptoverticillium, and Streptomyces.
- 10 19. The method of claim 17 wherein said genus is selected from the group consisting of Saccharapolyspora and Streptomyces.
 - 20. The method of claim 19 wherein said genus is *Saccharapolyspora* and the species is *erythraea*.
 - 21. The method of claim 19 wherein said genus is *Streptomyces* and the species is *hygroscopicus*.
- 22. The method of claim 1 wherein said polyketide is selected from the20 group consisting of macrolides, tetracyclines, polyethers, polyenes,ansaymcins and derivatives or analogs thereof.
 - 23. The method of claim 22 wherein said polyketide is a macrolide.
- 25 24. The method of claim 23 wherein said macrolide is an erythromycin.
 - 25. The method of claim 24 wherein said erythromycin analog is selected from the group consisting of 11-oxo-11-deoxyerythromycin A, 7-hydroxyerythromycin A, 6-deoxy-7-hydroxyerythromycin A, 7-
- 30 oxoerythromycin A, 3-oxo-3-deoxy-5-desosaminylerythronolide A, Δ-6,7-anhydroerythromycin A, ((14S,15S)14(1-hydroxyethyl)erythromycin A, 11-epifluoro-15-norerythromycinA, 14-(1-propyl)erythromycin A, 14(1-propyl)erythromycin A, and 14[1(1-hydroxypropyl)]erythromycin A.
- 3 5 26. The method of claim 1 wherein said DNA sequence, designated eryA, encodes the enzymatic activities associated with the formation of 6-deoxyerythronolide B.

PCT/US92/00427

3

- 27. The method of claim 26 wherein said DNA sequence comprises: the DNA sequence of Figure 2.
- 28. The method of claim 1 wherein said gene-containing DNA
 5 sequence encodes one or more enzymatic activities in the rapamycin biosynthetic pathway.
 - 29. The method of claim 23 wherein said macrolide is a rapamycin analog.

30. A compound selected from the group consisting of 7-hydroxyerythromycin A; 6-deoxy-7-hydroxyerythromycin A; 7-oxoerythromycin A, 3-oxo-3-deoxy-5-desosaminyl-erythronolide A; Δ-6,7-anhydroerythromycin A; ((14S,15S)14(1-hydroxyethyl)erythromycin A; 11-epifluoro-15-norerythromycin A; 14-(1-propyl)erythromycin A; 14(1-propyl)erythromycin A; and 14[1(1-hydroxypropyl)]erythromycin A.

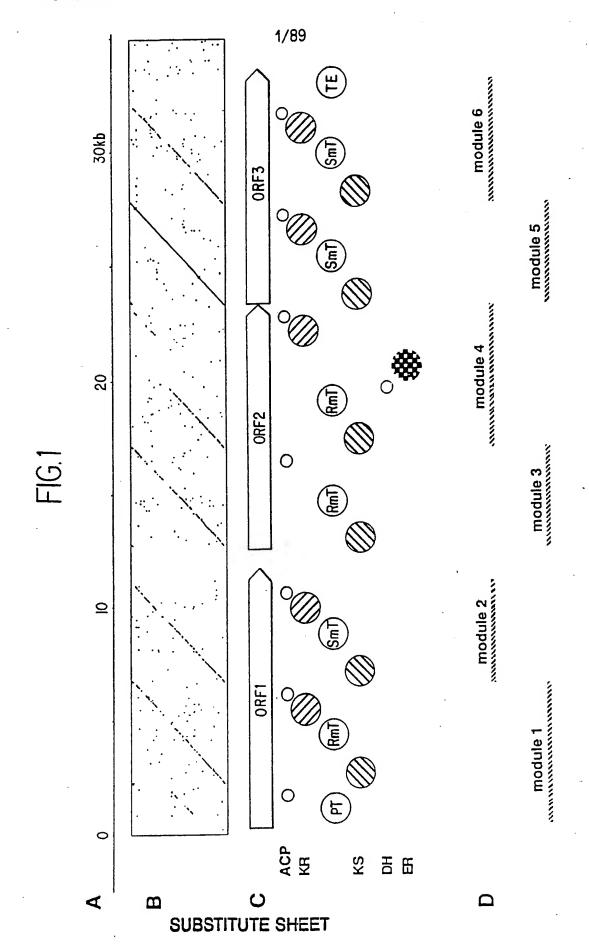


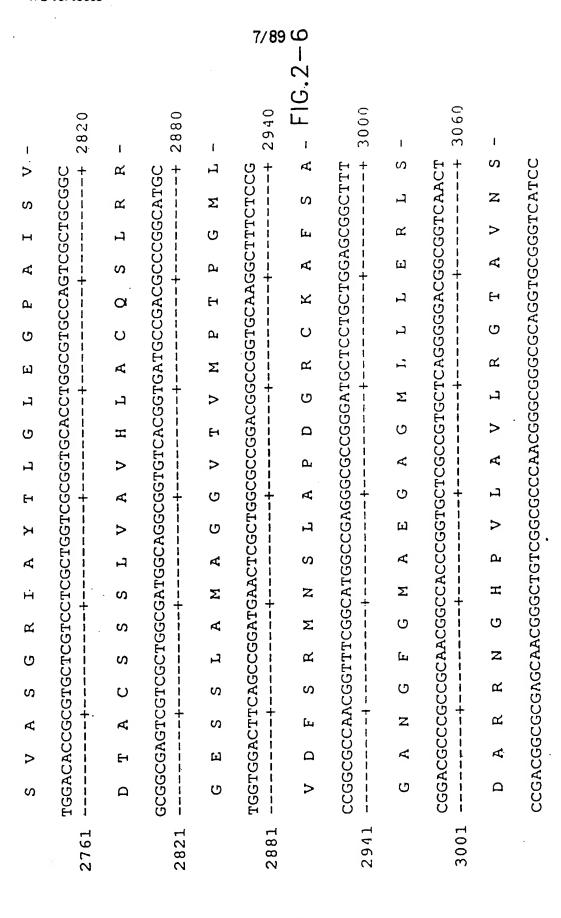
						FIG.2-					
í	1080	ı	1140	1	1200	ī.	1260	ı	1320	i	1380
L L G E S R V F A A A M D A C A R A F E	AGCCCGTGACCGACTGGACGCTGGCGCAGGTCCTGGACTCCCCGAGCAGTCGCGCCGCG	PVTDWTLAQVLDSPEQSRRV	TCGAGGTCGTCCAGCCCGCCCTGTTCGCGGTGCAGACGTCGCTGGCCGCGCTCTGGCCGCTTTGGCCGCTTTGGCCGCTTTGGCCGCTTTGGCGCTTTTGGCGCTTTTGGCGCTTTTGGCGCTTTTTT	EVVQPALFAVQTSLAALWRS	CCTTCGGCGTGACCCCCGACGCCGTGGTGGGCCACAGCATCGGCGAGCTGGCCGCGCGCCCCCCGCCCCCCCC	F G V T P D A V V G H S I G E L A A A H	ACGIGIGCGGIGCGGCCGGIGCCGCCGCCGCGCGCGCCCCCC	V C G A A G A A D A A R A A L W S R E	AGATGATTCCGTTGGTGGGCAACGGCGACATGGCAGCCGTCGCGCTCTCCGCCGACGAGA	MIPLVGNGDMAAVALSADEI	TCGAGCCGCGCATCGCCCGGTGGGACGACGTGGTGCTGGCCGGGGTCAACGGTCCGC 1321+++
							SŲE	BSTI	TUTE	SHE	ET

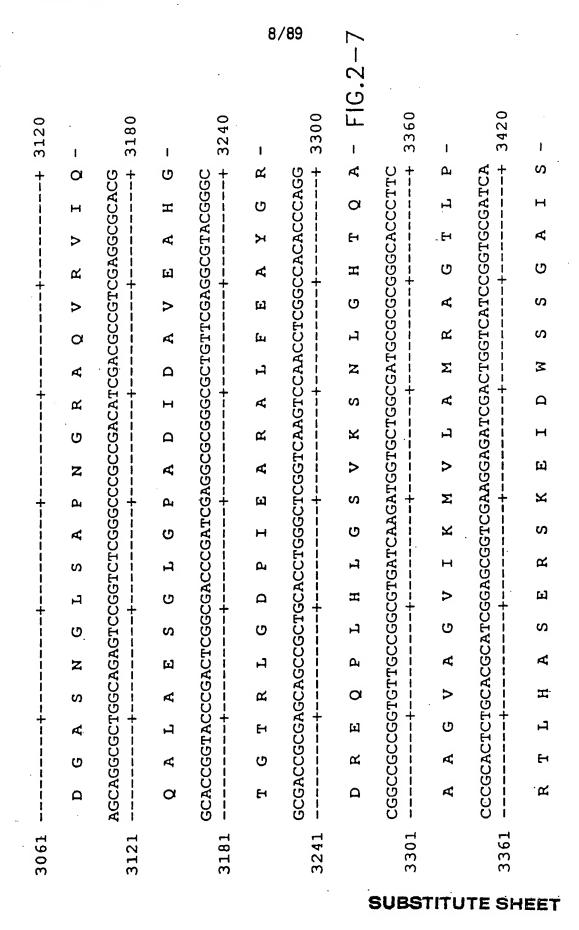
					3/89	-2				
						FIG.2-				
	1440		1500		1560	_ FIG	1620	ı	1680	ı
٦ ا		A A		- Q		>		×		ធា
വ	CICC	S	3GTC	>	GGA(ជេ	CGA	Ω	CCT	귝
ပ	SCTC	ч	3CAC	Ø	CHC	လ	000	æ	292	æ
z	3GAC	ធ	3GC(K	rgg(+	Ŋ	GGT +	>	GTC +	S
> .	CCAC	a	CTC	တ	000	Ö	GCT	ħ	555	ĸ
G	3GT(>	GCA	Ħ	225	Д	GGA	ធា	GAT	H
Ø	SCCGGAACCGGTCGCGCCCGGGTCCAGGAGCTCTCGG	K	GATGGCGGCGCACTCGGCGCAGGTCG	Ø	rcggcccrggcgrggtrcgcgcccggrggcrcggAgg	Æ	CTCACCGGAGGTGCGGTCGACACGCGGGAGCTGGTGGCCGACT	ሺ	CCGGCTGCCGGTGCGCTTCGACGAGGCGATCCGGTCCGCCCTGG	Æ
ъ.	+ 900	ĸ	† 1	æ	GTT 	ĹĿı	CAC	H	CGA	Ш
>	252	æ	GAT 	Σ	GTG:	3	, CG &	Ω	,CG2	Ω
>	GGT	>	GTC	တ	299,	A	3661	>	CTI	ĹĿ
Ω	ACC -+-	<u>ρ</u> .	ATGTG	>) - -	ы	3TGC	æ	rGC0	CL.
Ω	GGA	ជា	CAA	Z	0993	Æ	3AGC	Ŋ	3663	>
Ω	229	Ф	CAT	н	CIC	လ	5000	Ŋ	1GCC	Ф
3	GTC +-+	လ	+	>	TGCG	ፈ	CAC +	H	3GCJ -+	H
ĸ	0000	Ŋ	ACAC	ø	GAT	Σ.	3CC1	H	2001	ద
Æ,	GAC	H	9000	A	₹ 1	Ŋ	CCAC	S	SCTT	រុំក
H	GCI	H	200	<u>~</u>	CGCCGAG	ជា	ACGC	æ	3CAC	တ
ĸ	TCTGC	H	3667	>	10.00 1	4	rc T Z	×))) (CC.
щ		>	7660	Ŋ	ACAT	H	CLU	ᆈ	360	~
ᅜ	GCTCGGTTCTGCTGACCGGGTCGCCGGAACCGGTCGCGCGCCCGGGTCCAGGAGCTCTCGG	တ	CCGAGGGGGTCCGCGCACAGGTCATCAATGTGTCGATGGCGGCGCGCACTCGGCGCAGGTCG	ជា	ACGACATCGCCGAGGGGATGCGCTCGGCCCTGGCGTGGTTCGCGCCCGGTGGCTCGGAGG	Ω	TGCCCTTCTACGCCAGCCTCACCGGAGGTGCGGTCGACACGCGGGAGCTGGTGGCCGACT	а	ACTGGCGCCGCAGCTTCCGGCTGCCGGTGCGCTTCGACGAGGCGATCCGGTCCGCCCTGG	3
	1381		1441		1501		1561		1621	

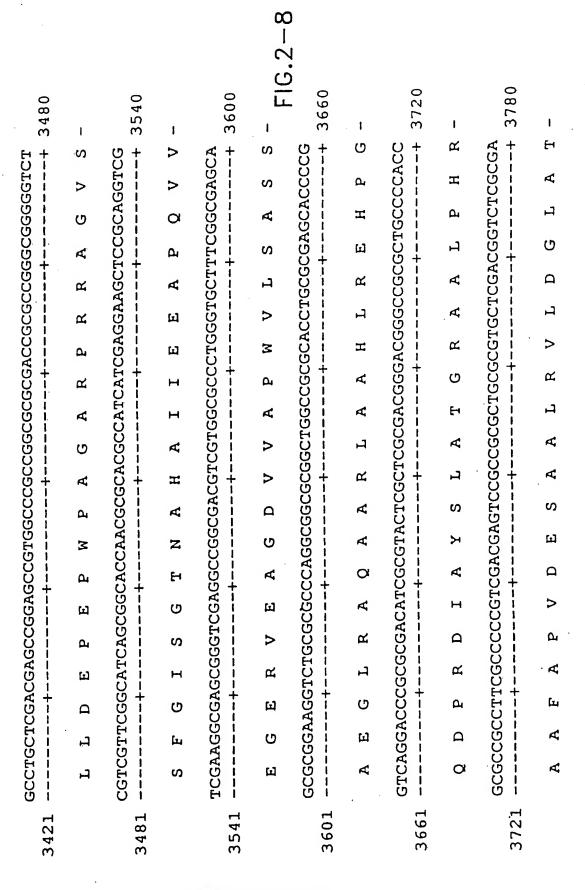
					(FIG.2- 20		,			
1740	i	1800	i	1860	i	F1(1	1980	 I	2040	1
AGGTCGGTCCCGGCACGTTCGTCGAAGCGAGCCCGCACCCGGTGCTGGCCGCCGCCGCGCTCC ++++++	V G P G T F V E A S P H P V L A A A L Q -	CAGACGCTCGACGCCGAGGGCTCCTCGGCCGCGGTGGTCCCGACGCTGCAACGCGGGC	QTLDAEGSSAAVVPTLQRGQ-	AGGGCGGCATGCGGCGGTTCCTGCTGGCCGCGCCCAGGCGTTCACCGGCGGCGTGGCCG		TCGACTGGACCGCCGCCTACGACGACGTGGGGCCGAACCCGGCTCTCTGCCGGAGTTCGC	DWTAAYDDVGPNPALCRSSR-	GCCGGCCGAGGAGGAGACGAGCCGGCCGAGTCCGGCGTCGACTGGAACGCGCCACCGCA ++	RPRRKTSRPSPASTGTRHRT-	CGTGCTGCGAGCGGCTGCTCGCGGTCGTCAACGGCGAGACCGCCGCGTTGGCGGGCCGCG	C C E R L L A V V N G E T A A L A G R E -
A 1681 -		AG 1741		A 1801 -		T 1861 -		1921 -		1981 -	

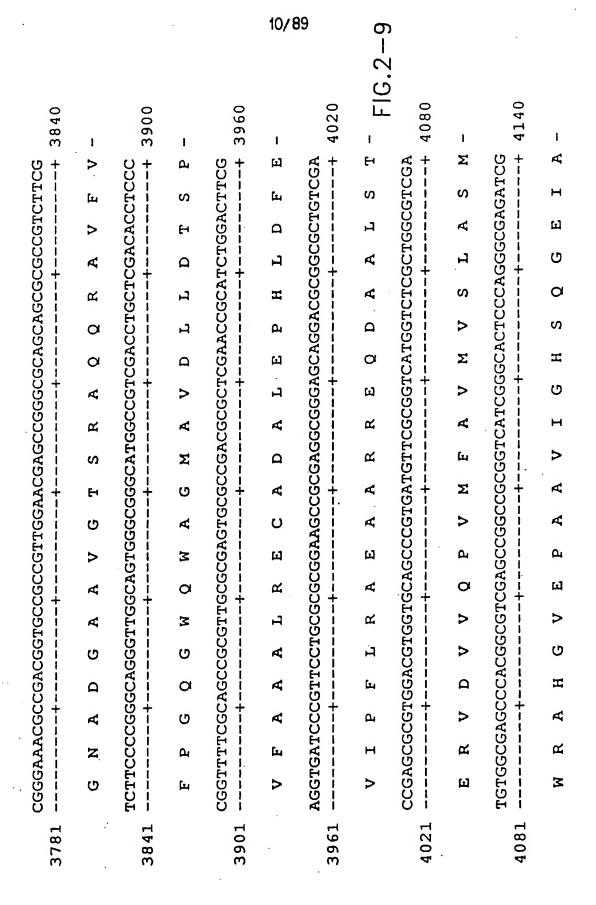
				5,	/89 -	F16.2-4				
2100	ı	2160	1	2220	ı	22	1	2340	ا ق	2400
AAGCCGACGCCGAGGCCACGTTCCGCGAGCTGGGGCTCGGTGGTGCTGGTGCCGCGCGCG	ADAEATFRELGLDSVLAAQL	TGCGCGCCAAGGTGAGCGCCGCGATCGGGCGCGAGGTCAACATCGCCCTGCTCTACGACC	RAKVSAAIGREVNIALLYDH	ACCCGACTCCGCGTGCGCTCGCGGAAGCACTCGCGGGGGGGAACCGAGGTCGCACAACGGG	PTPRALAEALAAGTEVAQRE	AAACCCGCGCGCGGACCAACGAAGCGGCGCCCGGCGAACCGGTCGCGGTCGTCGCGATGG	TRARTNEAAPGEPVAVAMA	CCTGCCGGCTGCCCGGCGTGTGAGCACCCCGGAGGAGTTCTGGGAGCTGCTGTCGGAGG	C R L P G G V S T P E E F W E L L S E (GCCGCGACGCGGTCGCGGGACTGCCGACCGACCGCGGCTGGGACCTGGACTCGCTGTTCC
2041		2101		2161		2221		22		23

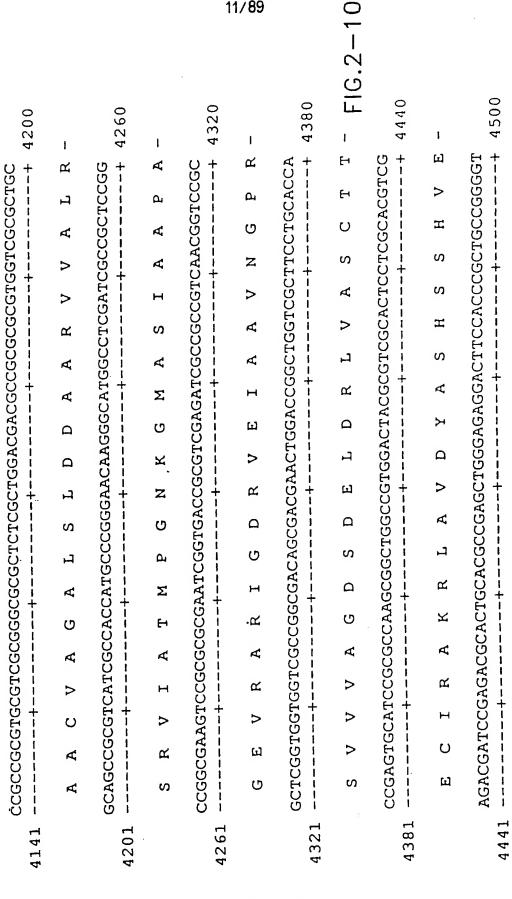
					6/8						
	-					FIG 2 -	1				
	2460		2520		2580		2640		2700	i	2760
H		ا «د		ا م		l Di		່ ບ		H	
Fri -	GAGG +	ы	SGACC	Ω	+	Д	TAC	×	ACCA	H	AGCG
J H)) 	H	STC(>	ATC.	н	GAG 	ធ	ACG 	H	ATC
S	CTG	1	3000	4	GGA	Ŋ	CAG	ø	GGT	Ŋ	909
Ω Ω	rtccr +	Г	CTGGC	H	CGAGCTCTCCTGGGAAGTGCTGGAACGGGCGGGGAATCCCGC	Æ	ATCCCGCAGGAGTACG ++	ы	GATGACCGGTA +	H	0000
ı	3GTJ	u ن	300	,	CGG(œ	ATC 	н	ATG	Σ	0999
Ω	995	ڻ ن	SAG(· ·	SAA	ធ	CTG	ı	CHI	1	GAG
	3CGGCG		+	<u>к</u>	7.TG(H	3 1 + 1	Ŋ	TAC -+-	×	CCTGG
3	909	_U	090		TGC	>	GTCGGCCT +	>	GGGCTACCT	ڻ ن	3GC
Ŋ	AGC	æ	CCC	Д	AAG		JTC(ניו	3AG(ப	CH C
K	ACC	a	TGT 	S	CCTGGG	<u>ロ</u>	GGCGTGTTC		STCC FFF		CTACACGCT
Ω	3CGC7	Ħ	GCATG	Σ	CCT	3	900	>	GGCGTC +	>	ACA
H	900	K	lCG(G	ICT	S	CTG	Ŋ	AAG	Ŋ	CCI
Д	CAC	H	CTO	ĮŦi	Jec.	1	CCA	H	908	[1].	TCG
Н	999	Ŋ	CTT	لتر		臼) dC(Д	3000	Ŋ	3CA -+-
ഗ	CIC	တ	099	K	ည်	Н	CTC	လ	3GG(ט	30
Ø	909	ĸ	000	Д	CAT	Σ	, GGC	A	75 CG	ធ	200
>	CAC	E	GCGTTCGA	Ω	GCAGCGGCTCATGCT	H	GC?	Ø	3660	L A	CCTC
æ	GACCCC	Д	GTT +	ĹĿij	929	ъ	CGTIG	ы	55.	H	3000
Ω	GA	Ω	000	K	SCA	à	CTC	လ	909	K	CG1
æ	ACCCCGACCCCACGCGCTCGGGCACCGCGCACCAGCGCGGCGGCGGGTTTCCTGACCGAGG	Δ	CGACCGCGTTCGACCCGGCCTTCTTCGGCATGTCCCCGCGCGAGGCGCTGGCCGTCGACC	Ħ	CGCAGCAGCGGCTCATGCT	a	CGACCTCGTTGCAGGCCTCGCCCACTGGCGTGTTCGTCGGCCTGATCCCGCAGGAGTACG	· H	GCCCGCGCTGGCCGAGGCGGCGAAGGCGTCGAGGGCTACCTGATGACCGGTACGACCA +++++++	. Д	CGAGCGTCGCCTCCGGCCGCATCGCCTACACGCTCGGCCTGGAGGGCCCGGCGATCAGCG
	2401		2461		2521				2641	•	2701









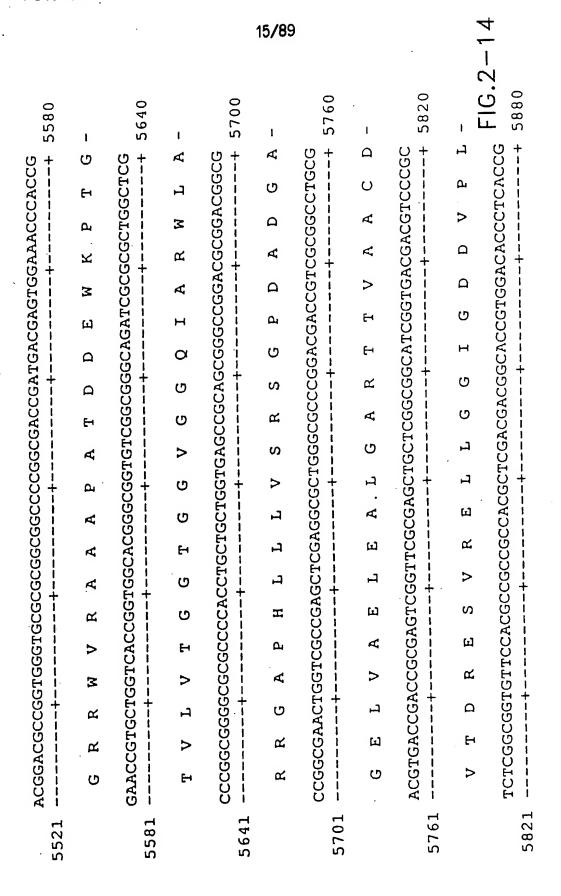


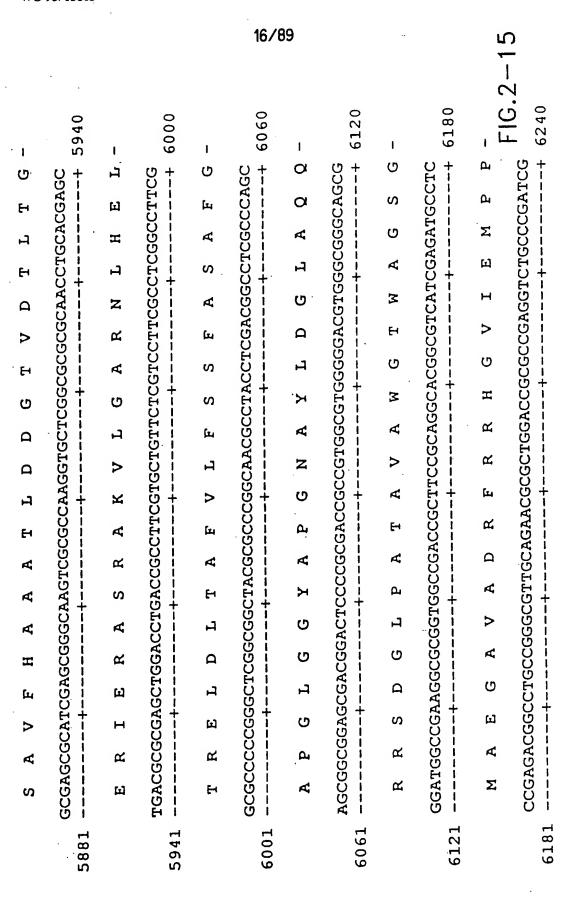
11/89

					12/89)		_	_		
						,		,	7-7		
	4560	1	4620	1	4680	1	4740	i	FIG.2-1	. 1	4860
TIRDALHAELGEDFHPLPGF	TCGTGCCCTTCTTCTCCACCGTCACCGGGCGCTGGACGCAGCCGGACGAGCTCGACGCCG	VPFFSTVTGRWTQPDELDAG-	GGTACTGGTACCGGAACCTGCGCCGCACCGTGCGGTTCGCGGACGCCGTCCGT	YWYRNLRRTVRFADAVRALA-	CCGAGCAGGGATATCGCACGTTCCTGGAGGTCAGCGCGCACCCGATCCTCACCGCCGCGA	EQGYRTFLEVSAHPILTAAI-	TCGAGGAGATCGGCGACGGATCGGGCGCCCGACCTCTCCGCCATCCAT	EEIGDGSGADLSAIHSLRRG-	GIGACGGCAGCCTCGCGGACTTCGGCGAAGCGCTCTCCCGGCGTTCGCCGCCGCGTGTCG	DGSLADFGEALSRAFAAGVA	CGGTGGACTGGGAGTCGGTGCACCTGGGCACCGGAGCACGCCGGGTGCCCTTGCCCACCT
	4501		4561		4621		4681		4741		4801

					13/89			7	-17	
ı	4920	I	4980	ı	5040	1	5100	(I	FIG.2-12 5160	i
×	ACCG	ជ	CGGTG	凹	CGAGA +	H	GGCCCTGGAGTCGGCCGGGCGCGGGTCCGCGAACTGGTCG ++	>	CGACGAACTCGCGGAGCGGCTTCGTTCGGTCGGCGAGGTGG +++++++	A
EH	CCAC	H		Q	CACCTGGCTGGCGAAGTACGCCGGAACCGCGGACGAGA +++	បា	TGG	>	AGG 	>
а	3GT(S	CGAGTGGCGGCCCACCGGTGC	A	2002	Ω	AAC	h	929	ជា
Н	+	K) +-	Ö	ACCGC	K	3CG	កោ	FCG -+-	Ŋ
Δ	929	ĸ	CAC	H	SAAC	H	2001	ц	7667	>
>	099	K	3600	Д	5000	Ŋ	3667	>	STTC	လ
ĸ	GGT	>	909	ĸ	GGCTGGTGGCGAAGTACGC	æ	000	ď	TCC	ĸ
æ	-+ 9229	д	GTG +	3	GTA	×	9999	Ø	3GCT	IJ
Ø	GAA	×	CGA	ធា	GAA	¥	090	ტ	2001	ď
Ŋ	ACC	Ы	CAT	Н	999	K	099	4	AACTCGCGGA	山
EH	CGA -+-	回	TACCG	R	GGT -+-	>	GTC	လ	-+-	Ø
Ŋ	GCT	ч	CTA	×	GCT	Н	GGA	ы	ACT	ч
ы	CTG	3	908 	ಜ	CCTG	3	GGCCCTGGAGTCGGC +	IJ	CGA	ធា
Ħ	CGT +	>	GCT +	П	CAC +	Ħ	+ 299	æ	CGACG	Q
>)) 	ĸ	0.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	Ø	 990	ပ	GGA 	ធ	TCG	ជ
တ	CGA(ជា	H H I	လ	CGA	Ω	TCG	æ	990	Ŋ
回	30.6(K	3GT	Þ	SCT	1	200	Æ	CTG	Ö
3	ccago	α	SGAG +-	ធ	+	ч	+ - + +	æ	CGCCCGCTGCG	ፈ
<u>Ω</u>	STT(Ľч	GA(Ω)))	A	CAC	H) 1 1 1 1	Æ
>	ACCCGTTCCAGCGCGAGCGCGTCTGGCTCGAACCGAAGCCGGTGGCGCGCCGGTCCACCG	Δı	AGGTCGACGAGGTTTCCGCGCTGCGCTACCGCATCGAGTGGCGGCCCACCGGTGCCGGTG	>	AACCCGCCCGGCTCGACGG	Д	CGAGCACCGCGGCTCGGGA	ဟ	TGGACGCCCGCTGCGGTCG	Ω
	4861		4921		4981		5041		5101	

					14/8	9				13	·
5220	1	5280	1	5340	ı	5400	ı	5460	i	FIG.2-15	1
CAGGAGTGCTGTCCCTGCTCGCGGTGGACGAAGCGGAGCCGGAGGAGGCGCCGCTCGCGC	GVLSLLAVDEAEPEEAPLAL	TGGCTTCGCTGGCGGACACGCTCAGCCTCGTGCAGGCGATGGTGTCGGCCGAACTCGGAT	ASLADTLSLVQAMVSAELGC	GTCCGCTGTGGACGGTGACGGAAAGCGCCGTCGCGACGGGGCCGTTCGAACGCGTCCGCA	PLWTVTESAVATGPFERVRN	ACGCCCCCACGCCCCTGTGGGGCGTCGGGCGGTCATCGCGCTGGAGAACCCCGCCG	AAHGALWGVGRVIALENPAV	TGTGGGGCGGCCTGGTCGACGTGCCCGCGGGGTCGGTCGCCGGAGCTGGCCCGGCACCTCG	WGGLVDVPAGSVAELARHLA	CGGCGGTCGTGTCCGGCGCGCGCGGTGAGGACCAGCTCGCGCTGCGCGCCGACGGGGTGT	AVVSGGAGEDQLALRADGVY
5161		5221		5281		5341		5401		5461	





-					17/8	19		(. . .		
•	0089	1	6360	i	6420	1	6480		FIG. 2-16	ı	
ETACRALONALDRAEVCPIV	TCATCGACGTCAGGTGGGACCGGTTCCTGCTCGCCTACACCGCGCAGCGCCCGGACCAGGC	IDVRWDRFLLAYTAQRPTRL	TCTTCGACGAGATCGACGACGCGGCGGGCTGCGCCGCAGGCGCCGGCCG	F D E I D D A R R A A P Q A P A E P R V	TGGGCGCGCTGGCGTCGCTGCCCGCGCGGAGCGCGAGGAAGCGCTGTTCGAGCTCGTGC	G A L A S L P A P E R E E A L F E L V R	GCTCGCACGCGGCCGCCGTCCTCGGCCACGCCTCGGCCGAGCGGGTGCCCGGCGGCCAGG	S H A A A V L G H A S A E R V P A D Q A	CCTTCGCGGAACTCGGCGTCGACTCGCTGTCGGCGCTTGAGCTGCGCAACCGGCTCGGCG	F A E L G V D S L S A L E L R N R L G A	CCGCGACCGGTGTCCGCCTGCCGACGACGTCTTCGACCACCCCGGACGTGCGGACGC
	T 6241 -		T 6301 -		1989		6421 -		6481		J

					18/	89			!	-1/	
0099	1	0999		0063		6780	ı I	6840	, i	FIG.2-1/	1
†	ч		æ	SCG	ן ט	+ 990	Æ	9 +	ڻ د	7CG	Ŋ
	H	ACCGG	Д	300	بط) 1 1 1 1	A	+	K	CTI	Įτι
	ĸ	3GC2	æ	3CT(ц	OH I	လ	000	A	GTT	Įъ
į	>	3600	æ	SCG	X	GCGA(Ω	TCCGA	·Q	00000	æ
Ī	Ω	GCAGGC	ø	TGC	ပ	ACG(~	CTC	S	000	æ
	Д	GAG	ធា	335	. 4	3662	Ŋ	360	A	CGA	Ω
	I	0001	A	CATC	Σ	CTC	S	GGGGCTGGGTCCCCGACGAGCTGATGGCCTCCGACGCGGCGG	Σ	CHH	Ĺij
+	Q	ACCGGAGC +	ပ	2992	ט	CACC	H	AGCTGA	ы	TGACT	Ω
	Ĺı	ACC	H	GTC	>	SATC	н	CGA(ធា	CGG.	ŗ
	>	000	A	ATC	н	SCIC	ı	CGA(Q		æ
+	H	CGGTG	<u>ა</u>	3600	- 4	GGA(ы	+-	Ф	+	Ŋ
	EH	0993	Ŋ	BATC	н	3TG(Z	3GT(>	3600	A
	EH	CTC	1	2225	Δı	SCTC	1	CTG	3	CAT(Σ
	بم.	GAP	ជ	GAC	回	3CG(ፈ	+-	Ŋ	ACTT(Ĺų
Ţ	ы	7900	æ	CGAC	. Δ	3GA(ជ		ĸ	CAA(z
į	r.	2000	æ	GTC	>	CCC	Сı	rga(Q	266	ტ
	>	CTC	1	CCC	Д	CTC	S	CGA.	Ω	CCA	H
+	Ŋ	SCAC	Ħ	3600	A P	GAC-+-	Ω	1000	Дı	-+	Æ
.	H	229:	æ	ACC	E	3GT(>	3GT(>)) - 	K
	Ø	TGGCGGCGCACCTGGCCGCCGAACTCGGCGGTGCGACCGGAGCCGAGCAGGCGGCGGCACCGG	Æ	CGACCACGGCCCCCGTCGACGAGCCGATCGCGATCGTCGCATGGCGTGCCGGCTGCCCG	₽	GGGAGGTCGACTCCCCGGAGCGGCTGTGGGAGCTGATCACCTCGGGACGCGACTCCGCGG	ធ	CGGAGGTCCCCGATGACC	四.	GAACCCGCGCCCACGGCAACTTCATGGCGGGCGCCGGTGACTTCGACGCGGCGTTCTTCG	E
6541		1099			7 0 0	6721		6781		6841	

.					19/	89				FIG.2-18	
0969	1	7020	1	7080	1	7140	ı	7200	. , i	F1G.	ı
GGATCTCGCCGCGCGCGCTGGCGATGGACCCGCAGCGCCCAGGCGCGCTGGAGACGA 1++++++	I S P R E A L A M D P Q Q R Q A L E T T	CGTGGGAGGCGCTGGAAAGCGCGGGCATCCCACCGGAGACGTTGCGCGGCAGCGACACCG	WEALESAGIPPETLRGSDTG	GCGTGTTCGTCGGCATGTCCCACCAGGGCTACGCGACCGGGCGTCCGCGCCCCGGAGGACG	V F V G M S H Q G Y A T G R P R D G	GCGTCGACGGGTACCTGCTCACCGGCAACACCGCGAGCGTCGCGTCGGGACGCATCGCCT	V D G Y L L T G N T A S V A S G R I A Y	ACGTGCTGGGGCTGGAAGGTCCCGCGCTGACGGTGGACACGGCGTGTTCGTCGTCGTTGG	V L G L E G P A L T V D T A C S S L V	TGGCGTTGCACACGGCGTGTGGGTCGTTGCGTGACGGTGACTGCGGTCTTGCGGTGGCCG	ALHTACGSLRDGDCGLAVAG
6901		6961		7021		7081		7141		7201	

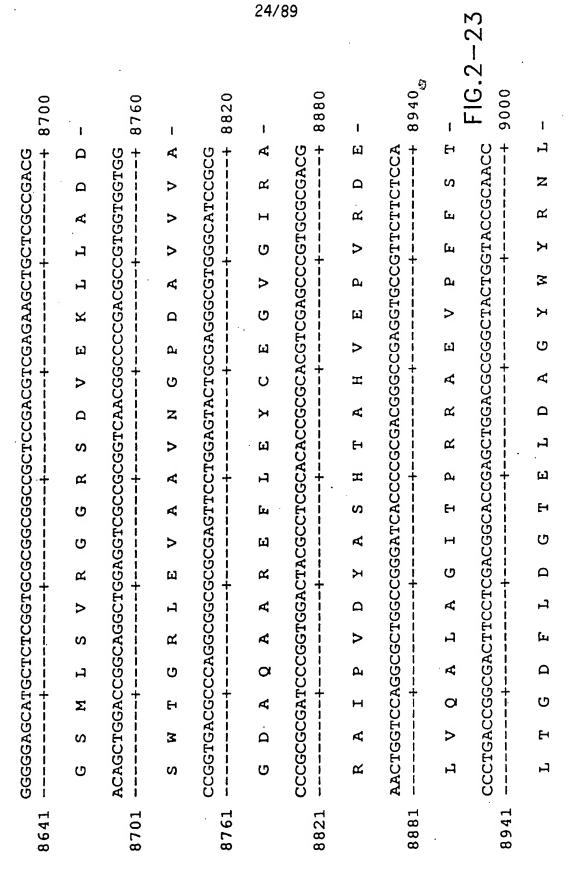
•			207	05			0	<u> </u>	
							C		
320	380	440		500		1560	- 10	<mark>۲ ا </mark>	
GTGGTGTGTCGGTGGGGGTCCGGAGGTGTTCACCGAGTTCTCCCGCCAGGGCGCGCCCCCCCC	acggccggtgcaagcccttctcggacgaggccgacggattcggtctcggg +++++	S P D G R C K P F S D E A D G F G L G E - AGGGTTCGGCGTCGTCCTCCAGCGGTTGTCCGACGCCAGGCGGGGGCCGCCGCG 	AFVULQRLSDARREGRRV-	rgctcggcgtggtggccgggtccgcggtgaaccaggacggcgcgagcaacgggctctccg +++++++	VVARGSAVNQDGASNGLSA-	CTCCGAGCGGCGTCGCGCAGCAGCGGGTCATCCGCCGGGCGTGGGCGCGTGCGGGGATCA	GVAQQRVIRRAWARAGIT-	cggcgcgcggatgtggccgtggtggaggcgcatgggaccggtacgcggctgggcgatccgg	DVAVVEAHGTGTRLGDPV-
GTGGTG	CTCGC	S P	S S	rGCTCG	L G	CTCCGA	P S	529995	G B
G 7261 -	T 7321 -	A 7381 -		T 7441 -	•.	C 7501 -		75.61	

20/89

					21	,	FIG.2-20			
7680	ı	7740	1	7800	1	7860	1	7920	(1	F1 G.
<pre>regaggcgrcgccrrgcraccractracgccaagrcgccgcgcrcgrcgcccgcrgc</pre>	EASALLATYGKSRGSSGPVL	TGCTGGGTTCGGTGAAGTCGAACATCGGTCACGCGCGGCGGCCGCGGGGTGTCGCGGGCG	LGSVKSNIGHAQAAAGVAGV	TGATCAAGGTGCTGCTCGGCCTGGAACGCGGTGTGGTGCCCCCCGATGCTGTGCCGGGGCG	I K V L L G L E R G V V P P M L C R G E	AGAGGTCGGGCCTCATCGACTGGTCCTCCGGCGAGATCGAGCTCGCAGACGCGCGTGCGGG	R S G L I D W S S G E I E L A D G V R E	AGTGGTCGCCCGCGGACGGGGTGCGCCGGGCAGGTGTGTCGGCGTTCGGGGTGAGCG	W S P A D G V R R A G V S A F G V S G	GGACGAACGCGCACGTGATCATCGCCGAGCCGCCGGAACCGGAGCCCGTGCCGCAACCGC
7621	-	7681		7741		7801		7861		7921

					2	2/89					F1 G. 2-21 340
	8040	1	8100	1	8160	ı	8220	1	8280	i I	F16.
NAHVIIAEPPEPEPVPQPR.	GACGCATGCTGCCCGCGACCGGGGTGGTGCCGGTCGTGCTGTCGGCCAGGACCGGGGCGG	M L P A T G V V P V V L S A R T G A A	CGTTGCGGGCGCAGGCCGGCAGGCTCGCCGACCACCTCGCCGCGCATCCCGGGATCGCAC	RAQAGRLADHLAAHPGIAP	CGGCCGACGTGAGCTGGACGATGGCGCGGGCCCGCCAGCACTTCGAGGAGCGGGCCGCGG	D V S W T M A R A R Q H F E E R A A V	TGCTCGCCGCCGACACCGCCGAGGCCGTGCACCGGTTGCGGGCGG	LAADTAEAVHRLRAVADGAV	TGGTTCCCGGTGTTGTCACCGGCAGTGCCTCCGACGGTGGTTCAGTGTTCGTCTTCCCTG	PGVVTGSASDGGSVFVFPG	GGCAGGGTGCCCAGTGGGAAGGCATGGCGGGGAGTTGTTGCCGGTTCCCGGTCTTGGCCG ++++++
,	GAC 7981	&	CGT 8041	IJ	CGC 8101	A	TGC 8161		TGC 8221	>	GG(8281

						-22				
ı	8400		8460	1	8520	ı	8580	ı	FIG.2-22 8640	1
O W E G M A R E L L P V P V F A E	CGCCGAGTGCGATGCGGTGTTGTCGGAGGTGGCCGGATTCTCGGTGTCCGAGG	ECDAVLSEVAGFSVSEV	TGCTGGAGCCACGTCCGGACGCGCCGTCGCTGGAGCGGGTCGACGTGGTGCAGCCGGTGC	R P D A P S L E R V D V V Q P V L	TGTTCGCGGTGATGGTGTCGCTGGCGCGGTTGTGGCGTGCCTGCGGTGCCGTTCCTTCGG	M V S L A R L W R A C G A V P S A	CCGTCATAGGGCACTCGCAGGGTGAGATCGCCGCCGCGGTGGTGGCGGGGAGCGTTGTCGC	H S Q G E I A A A V V A G A L S L	ATGCGCGTCGTCGCCCGCTCGAGGGCGGTGCGTGCGGTCGCGGGCC	M R V V A R R S R A V R A V A G R
0 G A C	agtcgatcgccgagtgcgat	SIA	TGCTGGAGCCACG	I E P	TGTTCGCGGTGA	E A V	CCGTCATAGGGC	D I	TGGAGGACGGCATGCGCGT	E D G
	8341		8401		8461		8521		8581	



					25/8	9				FIG.2-24
0906	ı	9120	ı	9180	ı	9240	1	9300	į	F1 G.
TGCGTCACCCGGTGGAGTTCCACTCCGCCGTGCAGGCGCTGACCGACC	RHPVEFHSAVQALTDQGYA,T	CGTTCATCGAGGTCAGCCCGCACCCGGTGCTGGCGTCGAGCGTCCAGGAGACCCTCGACG	FIEVSPHPVLASSVQETLDD	ACGCCGAGTCGGACGCGGCGGTGCTCGGGACGCTGGAACGCGACGCGGGCGACGCCGACC	AESDAAVLGTLERDAGDADR	GCTTCCTCACGGCACTCGCCGACGCGCACGCGCGGTGTCGCGGTCGACTGGGAAGCGG	F.L T A L A D A H T R G V A V D W E A V	TGCTCGGCCGGGCCGGACTGGTCGACCTGCCGGGTTATCCTTTCCAGGGCAAGCGGTTCT	LGRAGLVDLPGYPFQGKRFW	GGCTGCTGCCGGACCGCACCCCTCGTGACGAGCTCGACGGCTGGTTCTACCGGGTCG
9001		9061		9121		9181		9241		9301

					26,	/89				L	-75
1	9420		9480	ı	9540		0096	1	0996	, <u>L</u>	FIG.2-25 9720
Ω	+ -	>	ტ †	A	+ 900	>	+ 999	ា	5 + +	A	SCGT
>	3GT	>	CGAGG	ш	999	Ŋ	909	ሺ	CGA	Ω	000
ĸ	GTC	>) 1 1 1 1	Æ	000	æ	3GT	>	GGT	>	GGT
×	CTC	ı	CTC	H	ZTG(Ö	SCT(П	090	A	CCA
۲ı	rrggcr +	3	CGCGCT	æ	rgacto	Ω	TGTGCTG +	>	CGGCGCGGTCGACG	Ŋ	20000
Z	30.63	ĸ	TCC	S	3663		CCT	H	CTT	Ĺ	SCT
ග	3660	ŋ	AGGTGCGG +	æ	3GT(>	AAC(₽	CAC	H	366(
Ω	TGCGC	æ	GTG:	>	CTC	ı	ATGGTGCGGTGCAAAC +	α	GGTCA	>	GTGGG +
h	CTG	П	GAG	ഥ	999	Ŋ	GTO	>	ACC	Ħ	SCTC
ы	71GCCGCGC	A	GTC	>	999	_O	909	Æ	TGG	Z	AAG
Д	-	Ø	GGACCG +	EH	GTC +	>	GGT +	G	CCACTG	1	AGGCG
<u>س</u>	CCT	<u>م</u>	TGG	3	CGGCGTCGGCGGGCTGGT	ტ	GAT	Ω	CCA	പ	CAG
Д	GAA	ш	2993	ڻ ن	2525	æ	3660	G	9093	Æ	SGAC
E	TCC	S	GAC	Д	ACG	E	GAG	ជា	GAC	Ω	2000
E	+	æ	GAG	ы	GTC.	>	CCTC	1	ATC	H	2000
ፈ	+	а	GAG	凹	GAG	ធ	2000	A	3660	Ŋ	3600
Ω	GTG	>	CAC	H	000	Д	SCTO	r r	GAG	ഥ	3GTC
Д	CGAGGT	ធា	366GC	Ŋ	GAA	ഥ	TTG-	П	-+- :8cc	A	GCAGTCCGG
H	ACC 	E	GAG 	ធ	000	æ	TCG	တ	GAC	Q	AGI
н	ACTGGACCGAGGTGCCGCGCTCCGAACCTGCCGCGCTGCGCGGCCGTTGGCTCGTGGTGG	B	TGCCCGAGGGGCACGAGGACGGCTGGACCGTCGAGGTGCGGTCCGCGCTCGCCGAGG	Ωı	CCGGCGCCCGAACCGGAGGTCACGCGGGGGTCGGCGGGGCTGGTCGGTGACTGCGGGGCGGCGGGCG	ŋ	TGGTGTCGTTGCTCGCCCTCGAGGGCGATGGTGCGGTGC	>	AACTCGACGCCGAGGGCATCGACGCGCCACTGTGGACGGTCACCTTCGGCGCGCGGTCGACG	н	CGGGCAGTCCGGTGGCCCGCCCGGACCAGGCGAAGCTGTGGGGGCTGGGCCAGGTCGCGT
	9361		9421		9481		9541		9601	·	1996

					27/	89				(16.2–26
ı	9780	1	9840	ı	0066	ı	0966	ı	10020	, ī	
ഗ	99 +	ы	ტ + ე +	A.	ACG +	Ø	9 +	K	999	Ω)GG
æ	Ö	а	+	or.	3TA	×	SGT	>		М	GTC
>	CTGGACCGGCCTCGTCGACCTGCCGCACATGCCGGACCCGG	Ω		>	rccccccccacgrcaccccccccrccgagr	ធា	rccrggrcaccggrggcaccgccggccrgggcgcgggaggrgg -+++++++	ធា	GCCGAACACCTCGCGCTGGTCAGCAGGCGAGGCCCGG	Ŋ	GGT
a	D I	Д	797 +-+	Æ	CIC	S	090	æ	- +	ĸ	929
ပ	CATC	Σ	GGT(>	CAC	EH .	999	ഗ	CAG	ĸ	၁၅၁
1	SCA	Ħ	CCA	a	252	Ø	CCT	Н	CAG	လ	SSO,
Ŋ))) (ы	GGA(Ω	CAC	Ħ	990	ტ		>	GCT
3	CCF + - +	H	3GA 	ы	CGT 	>	+ 292	K	GCT +	Н	CCC
H	CGA(Ω	OH I	S	CCA	Ħ	CAC	H	292	æ	GAC
×	CGT	>	000	ტ	292	Æ	TGG	Ŋ	CCT	ы	GCT
A	CC F(П) i + :	A	-+-	Δ,	CGG -+-	Ŋ	ACA -+-	Ħ	CGA
a	000	ഗ	TGCTGGCCGGCTCGGAGGACCAGGTCGCGGT	П	TTC	S	CAC	E	CGA	Ш	2523
Ω	GAC	Н	GGT	>	GCT	H	199	>	090	Ø	GAC
ር _ብ	CTG +	3	+	æ	+-	ч	CCT +-+	П	20000	Ŋ	CCI
ĸ	CCG	ĸ		Ħ	909 	ជ	AAT 	н	TCG	ĸ	CGA
Ø	001	Д	TCT	7	TGC 	Æ	CAC	H	550	Ŋ	990
>	000	Ŋ	500	α,	909	œ	550	Ŋ	CGC	K	CGT
Д	ACGCG	CC.	GCGAGGCC	ტ	CGCCGTGC	>	+ 888	Ŋ	GCTC	ı	999
လ	3GA)	ជ	306	Ж	000	æ	0 1	Д	GTG	3	CGA
ტ	CCCTGGAACGCGGGCCCCG	ı	AACTGCGAGGCCGTCTCAC	IJ	CCGACGCCGTGCGTGCGCGGCGGCTTTCCCCCCGCCCACGTCACCGCCACCTCGGAGTACG	Ω	CGGTGCCGGGCGGCACAATCCTGGTCACCGGTGGCACCGCCGGCCTGGGCGCGCGGGGTGG	>	CCCGGTGGCTCGCCGGTCGCGGCGCCGAACACCTCGCGCTGGTCAGCAGGCGAGGCCCGG	æ	ACACCGAGGGCGTCGGCGACCTGACCGCCGGCTGACCCGGCTCGGCGCGCGC
	9721		9781		9841		9901		. 9961		

					;	28/89			i	-2/	
10080	1	10140	ı	10200	1	10260	1	10320	i	FIG.2-2/ 10380	ı
+	TEGVGDLTAELTRLGARVSV	TGCACGCGTGCGACGTCAGCAGCCGCGAACCGGTGAGGGAACTCGTGCACGGCCTGATCG	HACDVSSREPVRELVHGLIE	AGCAGGGCGACGTCGTCCGCGGTGTGGTGCACGCGGCGGGACTGCCGCAGCAGGTCGCGA	QGDVVRGVVHAAGLPQQVAI	TCAACGACATGGACGAGGCCGCCTTCGACGAGGTGGTCGCGGCCAAGGCCGGGGGGGCGGG	NDMDEAAFDEVVAAKAGGAV	TGCACCTGGACGAGCTGTGCTCGGACGCCGAGCTGTTCCTGCTGTTCTCCTCCGGGGCCGG	HLDELCSDAELFLLFSSGAG	GGGTGTGGGGGAAGCGCCCGCCAGGGCGCCTACGCCGCGGGCAACGCGTTCCTGGACGCCT	V W G S A R Q G A Y A A G N A F L D A F
10021		10081		10141		10201		10261		10321	

SUBSTITUTE SHEET

					2	9/89			. (-28	
10440		10500	1	10560	ı	10620	R. 1	10680	1	FIG.2-28	। ध
CCGCGGCCTGCCCGCCACGTCGGTGGCGTGGGGGGCTGTGGG	LPATSVAWGLW	CGACGAGGAGGCCGTGTCGTTCCTGCGCGAGCGCGGTGTGC +++++++	EAVSFLRERGV	CGCCCTCGCCGCCCTGGACAGGGTGCTGGCCTCCGGGGAGA	A A L D R V L A S G E	CGGCGGTGGTCGTGACGGACGTGGACTGGCCCGCCTTCGCCGAGTCCTACACCGCCGCCCC	WPAFAESYTAA	GGCCCCGGCCGTTGCTCGACCGCATCGTCACGACCGCGCCGAGCGAG	V T T A P S E R A G E	CGGAGACGGAGAGCCTGCGCGACCGGCTGGCGGGTCTGCCGCGTGCCGAGCGGACGGCGG	LAGLPRAERTA
TCGCCCGGCACCGCCGGGG	A'R H R R G R G	CGGCGGCGCCATGACCGG	AGGMTGDE	GGGCGATGCCCGTACCGCG	AMPVPRAL		AVVYTDVD		PRPLLDRI		E T E S L R D R
10381		10441		10501		10561		10621		10681	

			·		3	0/89			C	67-	
10800	i	10860	1	10920	:	10980	ſ	11040	(, ī	FIG.Z-29	i
AGCTGGTGCGCCTGGTCCGCACCAGCACCGCGACCGTGCTGGGCCACGACGACCGAAGG ++++++	STATVLGHDDPKA	CGGTGCGCGCGACCACGCCGTTCAAGGAGCTCGGGTTCGACTCGCTGGCGGCCGTCCGGC	KELGFDSLAAVRL	TGCGCAACCTGCTCAACGCGGCCACCGGGCTCCGCCTGCCGTCGACGCTGGTCTTCGACC	T G L R L P S T L V F D H	ACCCGAACGCCTCCGCGGTCGCCGGTTTCCTCGACGCCGAGCTCGGCACCGAGGTCCGGG	A G F L D A E L G T E V R G	TCGCCGGGCTGGACGCGCTGGAAGGCGCCCTGCCCGAGGTGC	A G L D A L E G A L P E V P	AGCTGGTACAGCGCTTGGAACGGATGCTCGCCGCGCGTACGCC	гуоягевмгаагке
	LVRLVRT	CGGTGCGCGCGACCACGC	VRATTPF		RNLLNAA		PNASAVA	GGGAGGCGCCGTCGGCCCTCGC	E A P S A L A	CCGCAACCGAGCGGGAAG	ATEREEL
10741		10801		10861		10921		10981		11041	

	1	31/89		30		
11160	11219	12702	12762	FIG.2-30	1	_
CGGTCGCCCAGGCCGCCGCCTCCGGGACCGGCGCCAACCCGGCGACGACCTGG 1++++++	GCGAGGCGGCGTGGACTGCTCGAAGCACTCGGCCGGGAGCTCGACGGCGATTGA	CCGCCGATTGGAGAAAAGGTGACTGACAGCGAGAGGTGGCGGAGTACCTCCGTCGGGCG	ACGCTCGACCTGCGTGCCGCCCGGCAGCGCATCCGCGAGCTGGAATCCGACCCGATCGCC	T L D L R A A R Q R I R E L E S D P I A ATCGTCAGCATGGCCTGCCGCGGGCGGGGTGAACACCCCGCAGCGGCTGTGGGAG 3	IVSMACRPORTER GGVNTPQRLWE	CTGCTGCGCGAGGGCGGTGAGACGCTGTCGGGCTTCCCCACCGACCG
11101	11161	12643	12703	12763		

						32/89				-	
۲۷		2		7		2		C1		FIG.2-31	
12882	i	12942	1	13002	ı	13062	ı	13122	1	FIG.	ı
	F P T D R G W D L	ACCCGGACAACCCCGGTACCAGCTACGTCGACAAGGGCGGG	GTSYVDKGG	GCTTCGACGCGGAGTTCTTCGGCGTCTCGCCGCGCGAGGCC	FFGVSPREA	AGCGGCTGCTGCTGGAGACGAGCTGGGAGCTGGTGGAGAAC	ETSWELVEN	GCCGGCATCGACCCGCACTCGCTGCGCGGTACCGCGACCGGCGTCTTCCTCGGAGTGGCG	ATGVFLGVA	ACACCGCCGCGGAGGACGTCGAGGGCTACTCGGTCACC	E D V E G Y S V T
+-	EGGETLSG	ACCACCCG	H H P D P D. N P	TTCCTCGACGACGCGGCGGCTTCGACGCGGAG	DAAGFDAE	GCGGCCATGGACCCGCAGCAGCGGCTGCTGC	1 D P Q Q R L L L	GCCGGCATCGACCCGCACTCGCTGCGCGGTACCG	OPHSLRGT	AAGTTCGCCTACGGCGAGGACACCGCCGCGGC	YGEDTAAA
12823	L L R	GCGCGGCTGC 12883+	A R L	TTCCTCG 12943	. F L D	GCGGCCA 13003	M M M	GCCGGCA 13063	A G I	AAGTTCG 13123	X Fr

	•	3	33/89	0	707
13242 -	13302	13362 -	13422	13482	F1 G. Z — S Z 13542 -
GGTGTGGCGCCCGCGGTCGCCTCCGGCCGCATCTCCTACACCATGGGCCTGGAGGGGCCG++++++-	TCGATCAGCGTCGACACCGCGTGCTCGTCGTCGCTGGTGGCCGCTGCACCTGGCGGTCGAG+++++++	TCGCTGCGCAAGGGCGAGTCGTCGATGGCGGTCGTCGCCGCGGTGCCGCGGTGATGGCGACC S L R K G E S S M A V V G G G A A V M A T	CCGGGGGTGTTCGTCGACTTCAGCCGGCAGCGCGCGCCGCCGCCGACGGGCGGTCGAAG ++++++	GCG	GAGCGGCTGTCGGAGGCGCGCGAAACGGGCACGAGGTGCTGGCGGTGGTTCGCGGCTCG 3+++++++
13183	13243	13303	13363	13423	13483

									1	3	
0.2		29		22		182		13842	c	FIG.2-5 13902	
13602	1	13662	1	13722	t	13782	ı	138	ı	13	11.5
GCGCTCAACCAGGACGGGGCCAGCAACGGGCTTTCCGCGCCGAGCGGGCCCGCGCGCG	ALNODGASNGLSAPSGPAOR-	AGGGTCATCCGGCAGGCCCTCGAGAGCTGCGGTCTGGAGCCCGGCGACGTCGACGCGGGGTG	RVIRQALESCGLEPGDVDAV	GAGGCGCACGGCACCGGTACGGCGCTCGGCGACCCGATCGAGGCGAACGCGCTGCTGGAC	EAHGTGTALGDPIEANALLD.	ACCTACGGCCGCGACCGCCGACCGGCCGCTCTGGCTGGGCTCGGTGAAGTCCAAC	TYGRDRDRPLWLGSVKSN	ATCGGCCACACCCAGGCGGCAGCGGCGTCACCGGCCTGCTGAAGGTGGTCCTGGCGCTG	I G H T Q A A A G V T G L L K V V L A L	CGCAACGGGGAACTGCCCGCGACCCTGCACGTCGAGGAGCCCCACGCCGCACGTCGACTGG	RNGELPATLHVEEPTPHVDW
13543		13603		13663		13723		13783		13843	

				35.	/89			7	-54 4
13962	14022		14082	1	14142	<u>.</u> 1	14202) L	FIG.Z-34 14262
CGTCCGGCGGCGTGGCGCTGCTGGCGGGCAAC	CGAATGCGCACGTGATCGTC	RRARVSAFGISGTNAHVIVE	GAAGCTCCTGAGCGCGCACCGGGAGACCACCGCGCACGACGGCCGACCGGTTCCGCTG	E A P E R E H'R E T'T A H D G R P V P L	GTGGTGTCCGCGCGCACGACGGCGGCGTTGCGGGCGCAGGCCGCCCAGATCGCCGAGCTG	V V S A R T T A A L R A Q A A Q I A E L	CTCGAACGCCCGGACGCCGACCTCGCCGGGGTCGGGCTGGGCCTGGCCACGACCCGCGCCC	LERPDADLAGVGLGLATTRA	CGCCACGAGCACCGCCGCCGTGGTGGCATCGACCCCGCGAGGAAGCGGTGCGCGGACTG
13		,	1 4		7,7		H		ਜ

					3	6/89				1	- 55 -
ŧ	14322	i	14382	1	14442	1	14502	1	14562	. (FIG. 2—55 14622
RHEHRAAVVASTREEAVRGL	CGGGAGATCGCCGCCGGTGCCGCGACGGCCGACGCCGTGGTCGAGGGCGTCACCGAGGTG	REIAAGAATADAVVEGVTEV	GACGGGCGCAACGTCGTCCTGTTCCCGGGGCAGGGTTCGCAATGGGCCGGCATGGGT	DGRNVVFLFPGQGSQWAGMG	GCCGAGCTGCTGTCGTCGCCGGTGTTCGCCGGGAAGATCCGGGCCTGCGACGAGTCG	AELLSSPVFAGKIRACDES	ATGGCCCCGATGCAGGACTGGAAGGTCTCCGACGTGCTGCGTCAGGCGCCCGGGGGCGCCCG	MAPMQDWKVSDVLRQAPGAP	GGCCTGGACCGGGTCGACGTGCTGCAGCCGGTGTTGTTCGCGGGGTGATGGTGTCGCTGGCG ++++++	G L D R V D V V Q P V L F A V M V S L A	GAGCTGTGGCGCTCGTACGGCGTGGAGCCCGCGGCGGTCGTGGGGGCACTCGCAGGGCGAG
	14263		14323		14383		14443		14503		14563

					37/	89			-36	
ı	14682	. 1	14742	ı	14802	ı	14862	I	FIG.2-36	ı
ធ ១ ០	ATCGCCGCCGCGCACGTCGCCGGGGGCGCTCACGTTGGAGGACGCGGGGAAGCTCGTCGTG	r v	GGCCGCAGCCGCCTGATGCGGTCGCTCTCCGGGGAGGGCGGCATGGCCGCCGCCGTCGCGCTG	V A L	GGCGAGGCCGCGGTGCGCGAGCGCCTGCGGCCGTGGCAGGACCGGCTCTCGGTGGCCGCG	V A A	TCGTGGTCTCCGGCGAGCCCGGCGCGCTGCGGGCGTTTTCC	A FI S	GAGGACTGCGCGGCCGAGGGCATCCGCGTCCGCGACATCGACGTGGACTACGCCTCGCAC	A S H
S	3CGAA	A K	GCTCTCCGGGGAGGGCGGCATGGCCGCCG	A	CTCTC	r S	CTGC	J R	GGACTACG	Z Q
G.	09090	A	CATG	Σ	00000	æ	 CCCG	Ą	CGTG	>
>	GGAGGA(+	Ω	3066	ტ	AGGA(- Δ	+	G	ACATCGA	Ω
>	TGGZ	ធ	3AGG(<u>ы</u>	rggc	Ø	SAGC	Е	GACA 	I Q
A A	ACGTT	T	09993	r U	3000	Ы		_ ნ	TCCGCGTCCGC	æ
Д	CGCTCAC	ı	CTCC	S	1GCG(ሺ	rcrc.	တ	SCGT	>
ធា	2999	A	CGCT	ᄓ	GCCT	H	TGG	>	TCC	H R
ک د	GCCG	Ø Æ	GGCCGCAGCCGCCTGATGCGGTC	R S	GAGC	ਜ਼ ਨ	GTCG	>	GGCA	<u>ن</u>
×	ATCGCCGCGCGCACGTCGC	>	GATG	Σ	GGCGAGGCCGCGGTGCGCG3	æ	GTCAACGGTCCCCGGTCGG'	တ	GGCCGAGG	ជា
S	GCA(Ħ	SCCT	'n	CGGT	>		æ	2992	æ
K	+ ++	A))))	æ	+	N N	GTC(G P	-+	C
L W	9008	A	CGCA	R	GAGG	д	GTCAACGGTCCCCG	z	GAGGACTGCGC	۵
ы	ATC	H	299	ტ		ڻ		>		ы
	14623		14683		14743		14803		14863	•

									1	(5-	
14982	ı	15042	1	15102	1	15162	1	15222	((FIG.2-5/	ı
TCGCCGCAGATCGAGCGGGTCCGCGAGGAACTCCTCGAAACGACCGGCGACATCGCGCCGG	SPQIERVŘEELLETTGDIAP	CGCCCGGCGCGCGGTGACGTTCCACTCCACTGTGGAGTCGCGGTCTATGGACGGCACCGAG	RPARVTFHSTVESRSMDGTE	CIGGAIGCCCGGIACIGGIACCGCAACCIGCGCGAGACGGIGCGCIICGCCGACGCCGIG	LDARYWYRNLRETVRFADAV	ACGCGGCTGGCGGAGTCGGGATACGACGCGTTCATCGAGGTCAGCCCGCATCCGGTCGTG	TR'LAESGYDAFIEVSPHP·VV	GTCCAGGCCGTCGAGGCGGTCGAAGAGGCTGACGGTGCCGAAGACGCGGGTCGTAGTC	V Q A V E E A D G A E D A V V	GGCTCGCTGCACCGCGACGGCGGTGACCTCTCGGCCTTCCTGCGGTCGATGGCCACCGCG	G S L H R D G G D L S A F L R S M A T A
14923		14983		15043		15103		15163		15223	

						. ••			1	FIG.2-38
15342	i	15402	i	15462	i	15522	1	15582	i	-
CACGTGTCCGGTGTGGACATCAGGTGGGACGTCGCTCTGCCCGGCGCGCGC	HVSGVDIRWDVALPGAAFFA	CIGCCGACGTATCCGTTCCAGCGCAAGCGCTACTGGCTCCAGCCCGCCGCACCCGCCGCCGCCGCCGCCGCCGCCGCC	LPTYPFQRKRYWLQPAAPAA	GCCTCCGACGAGCTGGCCTACCGCGTTTCCTGGACTCCGATCGAAAAGCCGGAGTCGGAA	ASDELAYRVSWTPIEKPESG	AACCTGGACGGCGACTGGTTGTCACACCCCTCATCAGTCCGGAGTGGACGGAAATG	N L D G D W L V V T P L I S P E W T E M	CTGTGCGAGGCCATCAACGCCAACGGTGGCAGGGCGTTGCGCTGCGAGGTGGACACGTCC	LCEAINANGGRALRCEVDTS	GCTTCGCGCACTGAGATGGCCCAGGCCGTCGCACAGGCCGGAACGGGATTCCGGGGCGTG 15583+++++
15283		15343		154		154	٠	15	-	15

39/89

					4	0/89				, ,	FIG. 2-39
ı	15702	1	15762	ı	15822	1	15882	i	15942	C L	F16.
ASRTEMAQAVAQAGTGFRGV	CTCTCGTTGCTGTCGTCGGACGAATCCGCCTGCCGTCCGGGGGGTTCCTGCCGGTGCGGTC	LSLLSSDESACRPGVPAGAV	GGCCTGCTCACCCTGGTCCAGGCGCTGGGCGATGCCGGGGTCGACGCACCGGTGTGGTGC 3+++++++	G L L T L V Q A L G D A G V D A P V W C	CIGACCCAGGGIGCGGICCGCACICCCGCCGACGACGACCICGCCCGGCCIGCGCAGACC	LTQGAVRTPADDDLARPAQT	ACCGCGCACGGCTTCGCGCAGGTCGCCGGGCTGGAGCTGCCGGGCCGCTGGGGCGGCGGTGTG	TAHGFAQVAGLELPGRWGGV	GTCGACCTGCCCGAATCGGTCGACGACGCGCGCGTGCGTCTGCTCGTGGCAGTCCTGCGC	V D L P E S V D D A A L R L L V A V L R	GGCGGCGGCCGTGCCGAGGACCACCTCGCGGTCCGGGACGGCCGCCTCCACGGCCGTCGC
	15643		15703		15763		15823		15883		15943
			0	HDC.	TITI ITE	CHE	ET				

ATCAGCGAAGTCGCACCGGAGGAGTTCGCCGAGACGATCGCGGCCAAGACCGCGTTGCTC

				•	41	/89			-40	
ı	16062	ı	16122	1	16182	1	16242	, L	F 16.2-40	ŧ
G G G R A E D H L A V R D G R L H G R R	GTCGTCCGCGCAAGCCTGCCGCAGTCCGGCTCGCGGAGCTGGACCCCGCACGGGACCGTG	V V R A S L P Q S G S R S W T P H G T V	CIGGICACCGGCGGCGAGCCCCGICGGCGACCAACIGGIGCGGTGGCTCGCCGACCGG	LVTGAASPVGDQLVRWLADR	GGAGCCGAGCGGCTGGTGCTGGCCGGAGCCTGTCCGGGCGACGACCTGCTGGCCGCGGTC	G A E R L V L A G A C P G D D L L A A V	GAGGAAGCGGGCGCATCGGCCGTCGTGTGCGCCCAGGACGCGGCGGCGCGCTGCGCGAGGCG	E E A G A S A V V C A Q D A A A L R E A	CTCGGCGACGAGCCGGTGACCGCGCTCGTGCACGCCGGAACCCTGACGAACTTCGGCAGC	LGDEPVTALVHAGTLTNFGS
	16003		16063		16123		16183		16243	

						42/89			•	-41	
16362	ı	16422	ı	16482	f,	16542	ı	16602	(1	FIG.2-	i
- ++++++	I S E V A P E E F A E T I A A K T A L L	GCCGTGCTGGACGAAGTCCTCGGCGACCGGGCCGTCGAGCGGGAGGTCTACTGCTCGTCG	AVLDEVLGDRAVEREVYCSS	GTCGCCGGGATCTGGGGCGCGCCGGGATGGCCGCCTACGCGGCAGGCA	VAGIWGGAGMAAYAAGSAYL	GACGCGCTGGCCGAGCACCACCGCGCGGGGCCGCTCGTGCACCTCGGTCGCCTGGACG	DALAEHHRARGRSCTSVAWT	CCGTGGGCGCTGCCGGGCGGGGGGGGTGGACGACGGCTACCTGCGGGAACGCGGGACTGCGC	PWALPGGAVDDGYLRERGLR	AGCCTCTCCGCCGACAGGGCGATGCGCACCTGGGAGCGGGTGCTGGCCGCCGCCGGGCCGGTG	SLSADRAMRTWERVLAAGPV
16303		16363		16423		16483		16543		16603	

	2		C 1	4	3 / 8 9		52		F1 G. Z-4Z	
16722	16782	ı	16842	ı	16902	ı	16962	Ĺ	17022	1
TCGGTCGCGGTGGCCGACGTGGACTGGCCGGTGCTCAGCGAAGGCTTCGCCGCCACC	CCGACCGCGCTGTTCGCCGAACTCGCCGGCGGCGGACAGGCGGAGGCCGAGCCGGAC	PTALFAELAGRGGQAEAEPD	AGCGGACCGACCGGCGAGCCGGCACAACGGCTCGCGGGGCTTTCCCCCGGACGAGCAGCAG	SGPTGEPAQRLAGLSPDEQQ	GAAAACCTGCTCGAACTCGTCGCGAACGCGGTTGCCGAGGTGCTTGGCCACGAGTCCGCC	ENLLELVANAVAEVLGHESA	GCCGAGATCAACGTGCGCCGCGCGTTCAGCGAGCTCGGACTCGACTCGCTCAACGCGATG	AEINVRRAFSELGLDSLNAM	GCCCTGCGCAAGCGCCTGTCGGCGAGCACCGGCCTGCGGCTGCCCGCGTCGCTGGTGTTC	ALRKRLSASTGLRLPASLVF
16663	. 16723		16783		16843		16903		16963	

					4	4 / 89			FIG 2-43) t	-
17082	1	17142	ı	17202		17262	ı	17322	- -	17382	
GACCACCCCACCGTCACCGCGCTCGCGCAGCACCTGCGCGCCCCGGCTCGTCGGTGACGCC	DHPTVTALAQHLRARLVGDA	GACCAGGCCGCGGTGCGCGTCGTCGGCGCGGCCGACGAGTCCGAGCCCATCGCCATCGTC 3	D Q A A V R V V G A A D E S E P I A I V	GGCATCGGCTGCCGTTTCCCCGGCGCCATCGGCTCGCCCGAGCAGTTGTGGCGGGTGCTG	GIGCRFPGGIGSPEQLWRVL	GCCGAGGGCGCGAACCTCACCACCGGCTTCCCGGCCGACCGGGGCTGGGACATCGGGCGG	AEGANLTTGFPADRGWDIGR	CTCTACCACCCGGACCCGGACAACCCCGGCACCAGCTACGTGGACAAGGGCGGGTTCCTC	LYHPDPDNPGTSYVDKGFL	ACCGACGCGGCGATTTCGACCCGGGCTTCTTCGGCATCACGCCCCGCGAAGCGCTGGCG	T D A A D F D P G F F G I T P R E A L A
17023		17083	SU	17143	ITE C	17203		17263		17323	
			30	BSTITL	IIE S	HEET					

					45.	/ 89			•	FIG.2-44
17442		17502	ı	17562	ı	17622	ı	17682	C L	17
ATGGACCCGCAGCACCCTCATGCTGGAGACGGCGTGGGAGGCAGTGGAACGCGGGGC	M D P Q Q R L M L E T'A W E A V E R A G	ATCGACCCCGACGCCCTGCGAGGCACCGGCGCGCGTCTTCGTCGGCATGAACGGCCAG	I D P D A L R G T D T G V F V G M N G Q	TCCTACATGCAGCTGCTGGCCGGTGAGGCCGAACGCGTCGACGGCTACCAGGGCCTCGGA	SYMQLL'AGEAERVDGYQGLG	AACTCCGCGAGCGTGCTCTCCGGGCGCATCGCCTACACCTTCGGCTGGGAGGGCCCGGCG	NSASVLSGRIAYTFGWEGPA	CTGACGGTGGACACCGCGTGCTCGTCCTCGCTGGTCGGCATCCACCTCGCGATGCAGGCG	LTVDTACSSLVGIHLAMQA	CIGCGGCGCGGTGAGIGCTCCCIGGCGCTGGCCGGCGGCGTCACGGTCATGTCCGACCCG
17383		17443		17503		17563		17623		17683

					46/	89		<u>ر</u> ر	 	
I	17802	ŧ	17862	ı	17922	ı	17982	- 6 0 7 4 5	18042	į
G V T V M S D P	TACACCTTCGTCGACTTCAGCACGCAGCGGGCTCGCCTCCGACGGTCGCTGCAAGGCG	A S D G R C K A	TTCTCCGCGCGGGCCGACGGCTTCGCGCTGTCGGAAGGCGTCGCCGCGCGCTGGTGCTGGAG	G V A A L V L E	CCGCTTTCCCGGGCGCGCGCCAACGGGCACCAGGTGCTGGCCGTGCTGCGCGGGGGGCGCGCG	L A V L R G S A	GCAACGGTCTCGCCGCTCCCAACGGCCCGTCGCAGGAGCGG	PNGPSOER	GTGATCCGGCAGGCGCTCGCCGCTTCGGGCGTGCCGGCCG	AADVDVVE
A L A G	GCAGCGCGGGCT	Q R G L	CGCGCTGTCGGA	A L S E	CGGGCACCAGGT	О Н Э	ACGGTCTCGCCGC	GLAA	CTTCGGGCGTGCC	S G V P
G E C S L	TACACCTTCGTCGACTTCAGCAC	V D F S T	TTCTCCGCGCGGGCCGACGGCTT	R A D G F	CCGCTTTCCCGGGCGCGCGCCAA	R A R A N	GTCAACCAGGACGGTGCCAGCA/	D G A S N	SCAGGCGCTCGCCG(O A L A A
L R R	TACACCITC	YTE	TTCTCCGCG	FS	CCGCTTTCC	P I	GTCAACCAC 17923+-	O z	GTGATCCGGCA	V I R

3

•	•				2	17/89			. (-46	
18102	1	18162	1	18222	I	18282	ı	18342	, <u>, , , , , , , , , , , , , , , , , , </u>	FIG.2-46	ı
GCGCACGGGACGGGCACCGAGCTCGGCGACCCGATCGAGGCCGGCGCGCTCATCGCGACC	AHGTGTELGDPIEAGALIAT	TACGGCCAGGACCGCGACCGGCCGCTGCGGCTCGGCTCG	Y G Q D R D R P L R L G S V K T N I G H	ACCCAGGCCGCGGCGCGCGCGGGGGTGATCAAGGTCGTGCTGGCGATGCGGCACGGG	T Q A A G A A G V I K V V L A M R H G	ATGCTGCCCCGGTCGTTGCACGCCGACGAGCTGTCCCCCGCACATCGACTGGGAGTCGGGG	M L P R S L H A D E L S P H I D W E S G	GCCGTGGAGGTGCTGCGCGAGGAGGTGCCGTGGCCGGCGGGTGAGCGCCCCCCGGCGGCG	AVEVLREEVPWPAGERPRA	GGGGTGTCGTCCTTCGGCGTCAGCGGAACCAACGCGCACGTGATCGTCGAAGAGGCACCA	G V S S F G V S G T N A H V I V E E A P
18043		18103		18163		18223		18283		18343	
			S	UBSTIT	UTE	SHEET					

·	• 2	48	3/89		747
18462 -	1.8522	18582	18642 -	18702	FIG.2-47 18762
GCAGAGCAGGACGCCGCACCGAGCGCGGTCCGCTGCCGTTCGTGCTGTCCGGCCGC+++++++	AGCGAAGCCGTGGTCGCGGCCCAGGCCCGCGCGCGCGCGC	GAGCTCGGCCTGACGCGGCGTGGACGCTCGCGACCGGCAGGGCGCGGGTTCGACGTG+++++++	CGAGCCGCCGTGCTCGGCGACGACCGCGGGCGTGTGCGCGGAGCTGGACGCGCTGGCC++++++	GAGGGCCGCCGTCGCCCGTCGCGCCGGTGACCTCCGCGCGCG	CTGGTCTTCCCCGGCCAGGGCGCGCAGTGGGCATGGCACGCGATCTGCTGGAATCC
18403	18463	18523	18583	18643	18703

					49/	89		·		FIG.2-48 9122	
18822	ı	18882	ı	18942	1	19002	ı	19062	i I	F16.	
TCCGAGGTGTTCGCCGAGTCGATGAGCCGGTGCGCCCGAGGCGCTCTCGCCGCACACCGAC	SEVFAESMSRCAEALSPHTD	TGGAAGTTGCTCGACGTCGTCGCGGCGACGGCGGTCCCGACCGCACGACGCGCGTCGAC	W K L L D V V R G D G G P D P H E R V D	GIGCICCAGCCGGIGCICTICICGAICATGGICTCGCIGGCCGAGCIGIGGCGCGCGCACACACACACACACAC	V L Q P V L F S I M V S L A E L W R A H	GGCGTGACCCCGGCCGCCGTCGGCCACTCGCAGGGCGAGATCGCCGCGGGCGACGTG	G V T P A A V V G H S Q G E I A A A H V	GCGGGCGCGCTGTCGCTGGAAGCCGCCGCGAAGGTGGTGGCCCTGCGCAGCCAGGTGTTG	AGALSLEAAAKVVALRSQVL	CGCGAGCTCGACGACCAGGGCGGCATGGTCGGTCGGCGCGCGTCCCGCGACGAGCTGGAG	
18		. 18		18		18	SŲI	STIT	UTE		

					50/	89				·	FIG.2—49
	19182	1	19242	ı	19302	1	19362	ı	19422	ر ا	19
RELDDQGGMVSVGASRDELE	ACCGTGCTCGCGCGCTGGGACGGCCGTGTCGCGGTGGCCGCCGTGAACGGGCCTGGCACC	T V L A R W D G R V A V A A V N G P G T	AGCGTCGTTGCCGGGCCGACCGCGGAGCTGGACGAGTTCTTCGCCGAGGCCGAGGCGCGGG	SVVAGPTAELDÖEFFAEAEAR	GAGATGAAGCCGCGCCGGATCGCCGTGCGCTACGCCTCCCACTCCCCGGAGGTGGCGCGCGC	EMKPRRIAVRYASPEVAR	ATCGAGGACCGGCTCGCGGCCGAGCTGGGCACCATCACCGCCGTGCGGGGCTCGGTGCCG	IEDRLAAELGTITAVRGSVP	CTGCACTCCACGGTGACCGGCGAGGTCATCGACACCTCCGCGATGGACGCCTCCTACTGG	LHSTVTGEVIDTSAMDASYW	TACCGCAACCTGCGCCGACCAGTGCTCTTCGAGCAGGCGGTGCGCGGTCTGGTCGAGCAG
	19123		19183		19243		19303		19363		19423
				UBS		SHE					

						517	-		-5(
					01		C:		FIG.2-50		
	19542		19602		19662		19722	,	F1G	ı	
1		l		I		ı		1			HC
α	+	ធា	TGCCGACGCTGCGCCGCGAGCAGAGC	လ	SCCGAC	Ω	SGGCCGGCCGAGCTGCCCACCTACCCGTTCGAACAC ++++	I	CGCTGGGCGTGCGCGGC	Ŋ	CTCGCCGCGGTCGACGTGCCGGGCCACGGCGGTGCTGTTC
ធ	CGA	ம்	AGC A	Ø	3000	Ø	ICG	ſΊ	TGC	K	SSS
>	199	>	3CG?	ជា		<u>ტ</u>	CGT	[±ı	908	>	GTG
ы	ATGGC	Æ	+-	æ	CGGCGTCGG	>	ACC(-+-	Δ,	TGG -+-	ڻ ء	909
O	rgat	X.	IGCO	æ	ACG(Ŋ	CCT	×	292	H	ACG
K	rgc1	H	2003	ħ	TGC	I	CCA	H	CGTCTCGG	A	SGCC
>	IGC:	П	CGA(+	H	CGGGCTCACG	>	TGC +	Δ,	TCT	S	555;
4	CGGTG	>	TGC(ሲ	CTC	=	AGC	1	ACG	>	TGC
α	ACC	Δ,	909	>	999	Æ	500	Ш	SCCGA	A D	BACG
ជ	CGC	I	CCT	O	TGC	ፈ	550	Æ	0000		STCC
ш	+	а	GTCACCT	E +	CCTGCTG	il.	+ :000:	<u>Д</u>	-+	я _Р	SCGC
ы	TGA	S)	AAG	>	ACC	7	SGAC	<u>α</u>	CACC	H	3000
>	AGG	>		田	GCA	z	3666	_ອ	2000	<u>а</u>	CICC
Δı	TCG	ы	+	4)TGC	J.	3000	4	-+	α H	CTG(
K	TCG	>		S G	GTTCCTG(F	3TG(>	+	<u>م</u>	CTG(
L R	CCT	ĮTI	ACG		SAG1	<u>ы</u>	3000	A.	TGG(3	SCG
	CGACAC	H	3AGC -+	н ы	CACO	H))) (ь 1	FEC.	Ĺ	CAC
z	TCG	Ω,	CGCCGAGCACGC		3000	4	CGT(ж -	GCGCTTCT	<u>۳</u>	GAG(
Y R	GGCTTCGACACCTTCGTCGAGGTGAGCCCGCACCCGGTGCTGCTGATGGCGGTCGAGGAG	G F	ACCGCCGAGCACGCGGGCGCGGAAGTCACCTGCGTGCCGACGCTGCGCCGCCGCGCGAGCAGAGC	T A	GGACCGCACGAGTTCCTGCGAACCTGCTGCGGGCTCACGTGCACGGCGTCGGCGTCGGCGCGTCGAC	S S	CTGCGTCCGGCGGTGGCCGGGGGGACGGCCGGCCGAGCTGCCCACCTACCCGTTCGAACAC +++++	L 1	CAGCGCTTCTGGCCGCGGCCGCCCGGCCCGACGTCTCGGCGCTGGGGCGTGCGCGGCG	α	GCGGAGCACCCGCTGCTG
~	19483 -	J	19543	•	19603		19663		19723		
	• •				·		<u> </u>	ıpe-	1 -1-4-1	= 011) pag-

						52 /	89		i	-51	
19842	I	19902	ı	19962	1 .	20022	1	20082	(1	FIG.2-51	ı
+++++	E H P L L L A A V D V P G H G G A V F	ACCGGAAGGCTTTCCACCGACGAGCAGCCGTGGCTGGCCGAACACGTCGTGGGCGGCCGG	GRLSTDEQPWLAEHVVGGR	ACGCTGGTGCCGGGCAGCGTCCTGGTCGATCTCGCGCTCGCCGCGGGTGAGGACGTCGGG	LVPGSVLVDLALAAGEDVG	CTGCCGGTCCTGGAGGAACTGGTGTTGCAACGGCCGCTGGTGCTGGCCGGGGCGGGGGGGG	PVLEELVLQRPLVLAGAGA	CTGCTGCGCATGTCGGTCGGCGCGCCCGACGAGTCGGGGCGGCGGACGATCGACGTCCAC	LRMSVGAPDESGRRTIDVH	GCCGCCGAAGACGTGGCCGACCTCGCCGACGCGCAGTGGTCGCAGCACGCCACCGGGACG	, A E D V A D L A D A Q W S Q H A T G T
19783 -	æ	A 19843 -	H	A 19903 -	H	19963	J	20023 -		20083 -	•

·						53/89			л С	F1G.2—J2 3502	
20202	ı	20262	ì	20322	ı	20382	1	20442	ر ر	7 1 6. 2	
.CTCGCGCAGGGCGTCGCGGGGTCCGAGGGATACCGAGCAGTGGCCGCCGCCGGAGGAC	LAQGVAAGPRDTEQWPPEDA	GTCCGCATCCCGCTCGACGACCACTACGACGGCCTCGCCGAGCAGGGCTACGAGTACGGA	VRIPLDDHYDGLAEQGYEYG	CCGTCGTTCCAGGCCCTGCGAGCCGCGTGGCGCAAGGACGACTCGGTCTACGCCGAGGTG	PSFQALRAAWRKDDSVYAEV	TCCATCGCGGCGGACGAGGAAGGTTACGCGTTCCACCCGGTGCTGCTCGACGCCGTGGCG	SIAADEEGYAFHPVLLDAVA	CAGACGCTCAGCCTGGGCGCCCTCGGCGAGCCGGGGGGGG	QTLSLGALGEPGGGKLPFAW	AACACCGTGACCCTGCACGCCTCCGGGGCGACCTCGGTGCGGGTCGTGGCGACGCCCGCC	NTVTLHASGATSVRVVATPA
20143		20203		20263		20323		20383		20443	

			·	54	/89			1	FIG.2-53
20562	20622	ı	20682	ı	20742	ı	20802	1	FIG.2
GGGGCGGACGCGATGGCCCTGCGGGTCACCGGCAGGCCACCTGGTCGCCACGGTC 03++++++++	GACTCGCTGGTCGTCCGCACCGGGGAGAAGTGGGAGCAGCCCGAACCGCGCGGTGGC	DSLVVRSTGEKWEQPEPRGG	GAGGGCGAGCTGCACGCTCTGGACTGGGGACGGCTAGCCGAGCCCGGCTCGACCGGTCGT	EGELHALDWGRLAEPGSTGR	GTGGTCGCGCCGATGCCTCGGACCTCGACGCCGTCCTGCGGTCCGGTGAACCCGAACCC	V V A A D A S D L D A V L R S G E P E P	GACGCGGTCCTGGTCCGCTACGAACCCGAAGGCGACGACCCCCGCGCGCG	DAVLVRYEPEGDDPRAARH	GGCGTCCTCTGGGCCGCCGCTCGTGCGCCGCTGGCTCGAACAGGAGGAGCTGCCGGGC
20503	20563		20623		20683		20743		20803
		CII	DOTITI	ITE S	HEET				

						5/89)				-54 -
1 -	20922		20982	ı	21042	1	21102	ı	21162	1	FIG.2-54
G V L W A A A L V R R W L E Q E E L P G	GCGACGCTGGTCATCGCCACGTCCGGCGCGGTCACCGTGTCCGACGACGACAGCGTTCCC	ATLVIATSGAVTVSDDDSVP	GAACCCGGCGCCGCCGCGATGTGGGGCGTGATCCGCTGTGCGCAGGCCGAGTCGCGGAC	EPGAAMWGVIRCAQAESPD	CGGTTCGTGCTCCTCGACACCGACGCGGAACCTGGGATGCTGCCTGC	R F V L L D T D A E P G M L P A V P D N	CCGCAGCTCGCGTTGCGCGGCGACGTCTTCGTGCCGCGCCTCTCGCGCGCG	PQLALRGDDVFVPRLSPLAP	TCCGCGCTGACGCTTCCGGCAGGCACCCAACGTCTCGTGCCGGGTGACGGGGGGGATCGAC	SALTLPAGTORLVPGDGAID	TCCGTGGCCTTCGAGCCCGCACCCGACGTCGAGCAGCCGCTCCGGGCGGG
	. 50		20	SUB	% STITUT	E SH			21		21
	-			نان							

						56/	89			1	-55
1	21282	. 1	21342	i	21402	i	21462	1	21522	i i	FIG.2-55
S V A F E P A P D V E Q P L R A G E V R	GIGGACGIGCGCCCACCGGAGICAACIICCGCGACGICCICCICGCACICGGCAIGIAI	V D V R A T G V N F R D V L L A L G M Y	CCGCAGAAGGCGGACATGGGCACCGAGGCCGCCGGTGTCGTCACGGCGGTCGGACCGGAC	P Q K A D M G T E A A G V V T A V G P D	GIGGACGCCTICGCGCCGGGAGACCGGGTGCTCGGCCTGTTCCAGGGAGCCTTCGCGCCG	V D A F A P G D R V L G L F Q G A F A P	ATCGCGGTCACCGATCACCGGCTCCTCGCACGAGTGCCGGACGGCTGGAGCGACGCCGAC	IAVTDHRLLARVPDGWSDAD	GCCGCGGCCGTGCCCATCGCCTACACCACGGCGCATTACGCGCGCTGCACGATCTCGCGGGG	A A A V P I A Y T T A H Y A L H D L A G	CTGCGCGCGGGTCAGTCGGTGCTCATCCACGCAGCGGCGGTGTCGGCATGGCGGCCCCCCCC
	21223	•	21283		21343		21403		21463		21523

					57/	/89			-56	
i	21642	ı	21702	ı	21762	ı	21822	l	FIG.2-	I
LRAGQSVLIHAAAGGVGMAA	GTCGCGCTGGCCCGCCGAGCGGGGGGGGGTGTTGGCCACCGCCGGCCCGGCCAAGCACACACA	VALARAGAEVLATAGPAKH	GGGACGCTGCGGCGCTCGGTCTCGACGACGAGCACCATCGCTTCCTCCCGGGAGACCGGT	G T L R A L G L D D E H I A S S R E T G	TTCGCCCGGAAGTTCCGGGAGCGCACCGGAGGCCGCGCGCG	FARKFRERTGGRGVDV'VLNS	CTCACCGGGGAACTGCTCGACGAGTCCGCGGATCTGCTCGCCGAGGACGGCGTCTTCGTC	LTGELLDESADLLAEDGVFV	GAGATGGGCAAGACCGACCTGCGGGACGCCGGGGACTTCCGGGGCCGATACGCCCCGTTC	EMGKTDLRDAGDFRGRYAPF
	21583		21643		21703		21763		21823	

				5	8/89				-57
21942	22002	1	22062	ı	22122	1	22182	1	FIG.2-57
GACCTCGGCGAGGCGGCTGACCGGCTCGGGGAGATCCTGCGCGAGGTCGTCGCCTG 3++++++	CTGGGCGCCGGGGAGCTCGACCGGCTCCCGGTATCGGCGTGGGAGCTGGGATCCGCGCCCC3++++++++	LGAGELDRLPVSAWELGSAP	GCGGCGTTGCAGCACATGAGCCGGGGCAGGCACGTCGGCAAGCTCGTGCTGACCCAGCCC 3++++++	AALQHMSRGRHVGKLVLTQP	GCGCCGGTGGACCCGGACGGCACGGTGCTGATCACGGGTGGCACCGGCACGCTCGGACGG	APVDPDGTVLITGGTLGR	CIGCICGCGCGCCACCICGICACCGAGCACGGCGIGCGGCACCIGCIGCTGGTCAGCAGG	LLARHLVTEHGVRHLLLLVSR	CGCGGCGCGGACGCGGGGTTCCGACGAGCTGCGCGGGAGATCGAGGACTTGGGCCGCG
21883	21943		22003		22063		22123		22183

						59	/89			1	- 58 -
1	22302	1	22362	ı	22422	1	22482	1	22542		FIG.2-
R G A D A P G S D E L R A E I E D L G A	TCCGCGGAGATCGCGGCTTGCGACACCGCCGACCGCGACGCGCTTTCGGCGCTGCTGGAC	SAEIAACDTADRDALSALLD	GGGCTGCCCCGGCCGCTGACCGGTGTCGTGCACGCGGGGGGGTGTGCTGGCCGGGCTGGCT	G L P R P L T G V V H A A G V L A D G L	GTCACCTCCATCGACGAGCCGGCGGTGGAGCAGGTGCTGCGCGCCAAGGTCGACGCGGCG	V T S I D E P A V E Q V L R A K V D A A	TGGAACCTGCACGAGCTGACCGCGAACACCGGTCTGAGCTTCTTCGTGCTGTTCTCGTCGTC	WNLHELTANTGLSFFVLFSS	GCGGCGTCGGTGCTAGCCGGCCCGGGGCAGGGCGTGTACGCGGCCGCGGCGGAACGAGTCGCTC	AASVLAGPGQGVYAAANESL	AACGCGCTGGCTGCCCTCCGGAGGACGCGGGCCTTCCCCGCGAAGGCGCTCGGATGGGGA
	. 22243		22303		22363		22423		22483	•	22543

						60/8	39			i	7-59
1	22662	1	22722	1	22782	ı	22842	1	22902	, (FIG.2-
NALAALRRTRGLPAKALGWG	CTGTGGGCGCAGGCCAGCGAGATGACCAGCGGACTCGGCGACCGCATCGCCCGGACCGGG	LWAQASEMTSGLGDRIARTG	GTCGCCGCGCTGCCGACCGAGCGGGCGCTCGCACTGTTCGACAGCGCCCTGCGCCGCGGC	VAALPTERALALFDSALRG	GGTGAGGTCGTGTTCCCGCTGTCCATCAACCGTTCCGCGCTGCGCAGGGCCGAGTTCGTG	G E V V F P L S I N R S A L R R A E F V	CCGGAGGTCCTGCGCGGCATGGTCAGGGCGAAGCTGCGCGCCGCCGGGCAGGCCGAGGCCG	PEVLRGMVRAKLRAAGOAEA	GÇAGGGCCGAACGTGGTCGACCGGCTCGCCGGTCCGAGTCCGACCAGGTCGCCGGG	AGPNVVDRLAGRSESDQVAG	CTGGCCGAACTGGTGCGTTCACACGCGGCGGCGGTCTCCGGGTACGGCTCGGCCGACCAG
	22603		22663		22723		22783		22843		22903

CGGGCGGACGCCCCGAGCACGTCCGCGATCAGCGAGGACGCCAGTGACGACGAGGTGTTC

					61/	89		(9 -	
									FIG.2-60 262	
	23022		23082	•	23142		23202	i	F1G 23262	•
ı	23	1	23	1		I		I		1
A A V S G . Y G S A D Q	CTCCCCGAGCGCAAGGCGTTCAAGGACCTCGGTTTCGACTCGCTGGCCGCGGGGGGGTGG ++++++	LGFDSLAAVEL	CGCAACCGCCTCGGTACCGCGACCGGCGTGCGGCTGCCCAGCACGTTGGTGTTCGACCAC	V R L P S T L V F D H	CCGACTCCGCTGGCGGTGGCCGAACACCTGCGGGACAGGCTGTTCGCGGCCTCTCACCGGCG 	LRDRLFAASPA	GTGGACATCGGCGACCGGCTGGACGAGCTGGAGAAGGCGCTCGAAGCCCTGTCCGCCGAG +++++++	LEKALEALSAE	GACGGGCACGACGTGGGCCAGCGCCTGGAGTCGCTGCTGCGCCGGTGGAACAGCAGG	LESLLRRWNSR
4	GAC	Д	3660	ტ	ACAC	I	CGAG	ជោ	3000	α.
H	CAAG	¥	GACC	H	CGA	ы	GGA(Ω	CCA(Ø
თ	GTT(+	Гц)))))))	A	IGGC -+	Æ	3GCT	니	+-	Ŋ
K.	AAGGCG1	Ø	TACC	Ħ	1993	>	-+	K.	ACGTG(>
>	CAA	X		ტ	299:	Ą	GTGGACATCGGCGA +	Ω	GACGGGCACGACGA	Ω
ч	CCGAGCGC	ĸ	CGCAACCGCCTC	H	ACTCCGCT	니	rcgo	ပ	ACG7	Ω
ы	CGA +	ជ	D))	K	JJEC -	Д	CA7	H	3GC7	I
æ	222	Д	CAA	z	GAC	H	662	Ω	090	Ŋ
ы	CTC	H		ĸ		۵		>		Ω
	22963		23023		23083		23143		23203	

	-				6	2/89			Č	FIG. Z—61 622	
23322		23382	1	23442	l	23502	I	23562	l L	F16.	ı
+ +	RADAPSTSAISEDASDDELF	TCGATGCTCGACCAGCGGTTCGGCGGGGGGAGGACCTGTAGATGAGCGGTGACAACGGC	SMLDQRFGGEDL*MSGDNG	ATGACCGAGGAAAAGCTCCGGCGCTACCTCAAGCGCACCGTCACCGAGCTCGACTCGGTG	M T E E K L R R Y L K R T V T E L D S V	ACCGCGCGCCTGCGTGAAGTCGAGCACCGGGCCGGTGAGCCGATCGCGATCGTCGGCATG	TARLREVEHRAGEPIAIVGM	GCGTGCCGGTTCCCCGGCGACGTGGACTCGCCGGAGTCGTTCTGGGAGTTCGTGTCCGGC	ACRFPGDVDSPESFWEFVSG	GGCGGGGACGCCATCGCGGAGGCCCCCGCCGACCGCGGCTGGGAGCCGGACCCCGACGCG	G G D A I A E A P A D R G W E P D P D A
23263		23323		23383		23443		23503		23563	

9 1	3GATCTCG + 23682 T S -
AGGCGCTGGCGATGGACCCGCAGCAGCGGATCATGCTGGAGATC	
PREALAMDPQQ	
GCGCTGGAGCGCCGCCGCCACGATCCGGTGTCCCTGCGCGGCAGCGCGACCGGGGTGTTC	GGGTGTTC + 23802
ALERAGHDPVSLRGSAT	63,
ACCGGTGTCGCCACCGTGGACTACGGCCCGCGACCCGACGAGGCCCCGGACGAGGTCCTG	23862
T G V G T V D Y G P R P D E A P D	E V L -
GGCTACGTCGGCACCGGCACCGCCTCCAGCGTCGCCTCCGGCCGG	ACTGCCTG
G Y V G T G T A S S V A S G R V A	г Х С Гг
GGCCTGGAAGGCCCGGCGATGACCGTCGACACCGCCTGTTCCTCCGGGCTCACCGCCTTG	ACCGCCCTG FIG.2-62
G L E G P A M T V D T A C S S G L	T A L

,					64/8	39			!	FIG.2-63
24042	1	24102	t	24162	1	24222	ı	24282	i i	7
CACCTGGCGATGGAGTCGCTGCGCCGGGACGAGTGCGGCCTGGCGCTGGCCGGCGGCGTG	H L A M E S L R R D E C G L A L A G G V	ACGGTGATGAGCAGTCCCGGGGCGTTCACCGAGTTCCGCAGGCCAGGGCGGGC	TVMSSPGAFTEFRSQGGLAA	GACGGCCGCTGCAAGCCGTTCTCGAAGGCCGCCGACGGGTTCGGCCTGGCCGAGGGTGCC	DGRCKPFSKAADGFGLAEGA	GGGGTCCTGGTGCTGCAACGGCTGTCGGCCGCGCGGCGGGGGGGCAGACCGGTGCTGGCC	G V L V L Q R L S A A R R E G R P V L A	GTGCTGCGGGGCTCGGCGGTCAACCAGGACGGCGCCAGCAACGGGCTGACCGGGCGCGAGC	V L R G S A V N Q D G A S N G L T A P S	GGACCCGCGCAGCAGCGGGTCATCCGCCGGGCGCTGGAGAACGCCGGTGTCCGGGCGGG
23983		24043		24103		24163		24223		24283

		٠			65	5/89					-64
ı	24402	1	24462	ı	24522	ı	24582	1	24642	1	FIG.2-64
GPAQQRVIRALENAGVRAG	GACGTCGACTACGTGGAGGCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGTG	D V D Y V E A H G T G T R L G D P I E V	CACGCGCTGCTCTCGACCTACGGCGCGGAACGCGACCCGGACGATCCACTGTGGATCGGT	HALLSTYGAERDPÖDPLWIG	TCGGTCAAGTCCAACATTGGCCACACCCAGGCCGCCGCCGCCGGCGTCGCCGGGGGTGATGAAG	SVKSNIGHTQAAAGVAGVMK	GCGGTGCTGGCGCTGCGGCGACGCGAGATGCCGCGCACGCTGCACTTCGACGAGCCCTCG	A V L A L R H G E M P R T L H F D E P S	CCGCAGATCGAGTGGGACCTGGGCGCGGTGTCGGTGGTGTCGCAGGCGCGCGGTGGTGCCCC	PQIEWDLGAVSVVSQARSWP	GCCGGCGAGAGGCCCCGCAGGGCGGCGTCTCCTCGTTCGGCATCAGCGGCACCAACGCG
	24343		24403		24463		24523		24583	·	24643

					66	6/89				L (O
							0.1		01	(FIG.Z-65 5062
	24762		24822	ī	24882	ı	24942	1	25002	1	F1G. 25062
1		1		•		·					99
æ	rcgggr	Ŋ	GGACGG +	ĸ	, dC	Н	ACGCC	æ	3GC +	α	AGGGG +
z	CTC	S	9993	Ŋ	tTCACGCTG	H	ACG.	Ω	CCGGGCAG	Ŋ	000
H	CACGTCATCGTCGAAGAGGCGCCCGAGGCCGACGAGCCCGAGCCGGCACCCGGACTCGGGT ++++++	Ω	CCGGTCCCGCTGGTGTTGTCCGGCCGCGACGAGCAGGCGATGCGGGCGCAGGCGGGCG	A		Įτι	TGGGAGCACCGCGCGGTGGTGGTCGGCGACCGCGACGACGCC	Ω	CA	H	GCCCGAACTCGCCGCGGCGTCGCGATGGTGTTCCCCCGGCCAGGGCGCGCGC
U	GGCACCC	Ω,	CGCA-	Q	ACACCGGT +	Ŋ	GACCG +	ĸ	GACGGC +	Æ	CCCGGCCAGGGCGCGCAGT
Ś	099	æ	9880	4	CAC	H	3CG7	Ω	3GA(H) 3 1 1
н	000	Д	929	ĸ	, CG	Ω	3000	Ŋ	ACCG	ĸ	AGG(
Ŋ	3AGCCCGA +	ы	GAT	A ~ M	300	ĸ	GGTGGT	> .	CGCCGAC	Ω	3CC7
Ĺι	0091	Д	+	Ø	CGTTGC +	7	TGG7	>	rcg	K	. i c
့တ	CGA	ា	GCA	a	CTO	S	3667	>	GCA	Н	ICC.
S	CGA 	Ω	CGA	ជា	GAP	z	3000	Æ	3000	К	IGT
>	3AGGCC	A	CGA -+-	Q	-+	ĸ	GCACCGCG	ĸ	SGCCGACGGCC	Ŋ	TGGT
Ŋ	CGA	ы	900	ĸ	0091	Д	AGC.	H		Ω	CGAC
æ	225	ф	990	Ŋ	CGA	ា	3667	Ы	1660	Æ	ICG(
ĸ	-	A	GTC	S); +	R.) -+	3	7667	>	GCG!
ಜ	AGAG	ធ	GTTG	н	200	K	3000	A	3000	A	3CG(
Д	CGA	ធា	GGT	>	CCI	1	SCAG	လ	rGCC	K.	2000
K.	CGT	>	GCT	Н	(C)	茁	3006	ĸ	3GCJ	н	CTC
凹	CATC	н	+ - CCC	വ	AGA	Ω	000	ជ	7000	Ŋ	SAA(
Ŋ	CGT	>	CCGGTCCCGCTGG	>	CTGGCAGACCACCTCGCC	Æ	GCCACCCGCCGCAGCGCG	E	CTCGCCGGGCTGCGCGCG	A	GCCCGAACTCG
K		Ħ		Д		ı		· 4		H	
	24703		24763		24823		24883		24943		25003

					67/89)		(99-	
1	25122	1	25182	í	25242	ı	25302	(((FIG.2-	i
ARTRGVAMVFPGQGAQWQG	ATGGCCCGCGACCTGCTGCGGGAGTCGCAGGTATTCGCCGACTCGATCCGCGACTGCGAG	MARDLLRESQVFADSIRDCE	CGGGCGCTGGCCCCCGCACGTCGACTGGTCGCTGACCGACC	RALAPHVDWSLTDLLSGARP	CTGGACCGGGTCGACGTCGTCCAGCCCGCGCTCTTCGCCGTCATGGTGTCGCTGGCGGCG	L D R V D V V Q P A L F A V M V S L A A	CTGTGGCGCTCCCACGGCGTCGAGCCCGCCGCGGTCGTCGGCCACTCGCAGGGCGAGATC	LWRSHGVEPAAVVGHSQGEI	GCCGCCGCGCACGTCGCCGCGCGCTCACCCTGGAGGACGCCCCCCAAGCTCGTCGCGGTC	A A H V A G A L T L E D A A K L V A V
	25063		25123		25183		25243		25303	

25422	25482	25542	25602	25(FIG.2-
CGGAGCCGGGTCCTGGCCCGGCTCGGCGCGCGCATGGCGTCGTTCGGGCTGGGC 25363+++++++-	ACCGAGCGGCCGAACGGATCGGGCGCTTCGCGGGCGCGCTCTCCATCGCCTCGGTC 25423+++++++	AACGGCCCCCGGTCGTCGTCGTCGCGGGGGAGCGGCCGCTGGACGAGCTGATCGCC 25483+++++++	GAGTGCGAGGCCATAACGGCGCGCCGCATCCCCGTCGACTACGCCTCCCACTCA 25543+++++++	CCGCAGGTGGAGTCGCTGCGCAGGAGCTGACCGAGCTGGCGGGCATCTCCCCGGTG 25603+++++++	TOGGCGGACGTGGCGCTCTACTCGACCACGGGCAGCCCATCGACACCGGCCACGATG T 25663

					69	9/89				. (ώ Σ
										ı	7-
	8.2		42		02	•	29		22		FIG.2-68
ı	25782	ı	25842	i	25902	ı	25962	1	26022	l	26(
Σ	+ ++	æ	CAGCTCGCCGAGGCGGGGTTCGACGCGTTCGTCGAGGTCAGCCCGCATCCGGTGCTGACC	H	GTCGGCATCGAGGCCACGCTGGACTCCGCGCTCCCGGCCGACGCCGGCGCGCCTGCGTG ++	>	GGCACCCTGCGGGGACCGCGGCGGCCTGGCCGACTTCCACACCGCGCTCGGCGAGGCG	æ	TACGCGCAGGGCGTGGAGGTCGACTGGAGCCCCCGCCTTCGCCGACGCGGGGGGGCGGTCGAG ++++	ப	CTGCCCGTCTACCCGTTCCAGCGGCAGCGGTACTGGCTGCCCATCCCCACCGGCGGGGCGC
E	GACACCGCCTACTGGTACGCGAACCTGCGCGAGCAGGTCCGCTTCCAGGACGCGACGCGG ++++++	Ħ	GCTG	H	CGTC	>	CGA(ា	CCGGT	>	
A	2900	Ø	3GT(>	CTG(ပ	000	Ŋ	9 9 1 1	Д	550
E	3GA(Ω	CCCGCATCCGG	Д	- +	K	GCT	ы	+	ĸ	CAC +
Ω	CCA(ø	SCA	x	000	Ŋ	CGC	Ø	252	Ø	000
н	CFF	[T-1	000	ф	292	Ø	CAC	Ħ	CGA	Ω	CAT
Д	GCAGGTCCGCTT +	ĸ	CGAGGTCAG	တ	CGA	Ω	CCA	Ħ	2002	Ø	000:
Ø	GGT +	>	GGT +	>	90099	Ø	CGGCCTGGCCGACTTCC	[F1	CTTC	[Li	;GC7
O	GCA	O	CGA	<u>,</u>	200	Δι	CGA	Ω	5555	æ	ACTO
H	CTGCGCGA +	ы	CGT	>	GCT	H	1660	A.	2225	വ	3GT?
H	929	ద	GCGTT	দ্য	- +	A	3CCJ	П	3GAGC +	လ	GCAGCGGT
H	CCI	Н)))	Ø	ACTO	ഗ	3000	Ŋ	ACT(Z	3GC?
S	GAA	Z	,CGP	Д	rgg2	Ω	3000	ტ	rcg.	Ω	AGC(
Ħ) +	Ø	.+	ក្រ	ACGCT	Ы	ACCC	æ	GGAGGT	>	FCC.
H	GTA	×	CGGGGTT	Ŋ	CCAC	H	3667	Ω	rgg.	ជ	E DC
æ	CTC	3	1660	Æ	4GGC	Ø	3600	ሺ	300	>	ACC(
>	CACCGCCTACTGGTACGC	×)CG?	ជ	CATCG?	ш	FGC0	М.	GCGCAGGGC	ტ	rcr;
Ē)CGC	A	rcgc	A	3CAT	н	2000 	H	7.090	Ø	CCG
æ	ACAC	H	CAGCTCGCCG	Ы	GTCG(ტ	GGCACCCTGCGGCGGGACCG +	E+	TACG	K	CTGCCCGTCTACCCGTTCCAG
S)		Ω		a		>		Ŋ		X	
	25723		25783		25843		25903		25963		26023

					70)/89				,	FIG.2-69
i	26142		26202	ì	26262	i	26322		26382	1	FIG. 2
LPVYPFQRYWLPIPTGGR	GCACGGGACGAGGACGACGACTGGCGCTACCAGGTCGTATGGCGGGAAGCCGAGTGGGAG	ARDEDDDWRYQVWREAEWE	AGCGCTTCGCTGGCCGGACGCGTGCTGCTGGTGACCGGACCGGGCGTGCCGTCCGAGTTG	SASLAGRULLVTGPGVPSEL	TCGGACGCCATCCGAAGTGGACTGGAGCAGAGCGGTGCGACGGTCCTGACCTGCGACGTG	SDAIRSGLEQSGATVLTCDV	GAATCCCGTTCGACCATCGGCACCGCACTGGAGGCCGCCGACACCGACGCTCTGTCCACT	ESRSTIGTALEAADTDALST	GIGGIGICGCIGCIGICCCGCGACGGCGAGGCCGTCGATCCGICGCTGGACGCGCTCGCC	V V S L L S R D G E A V D P S L D A L A	CIGGICCAGGCCCICGGAGCGGCCGGGGICGAAGCACCGCTGIGGGIGCTGACCCGCAAC
٠	26083		26143		26203		26263		26323		26383

	·				71/	'89	٠) 1 0	FIG.2-/0		
	26502	1	26562	1	26622	ı	26682	i I	267	1	
L V Q A L G A A G V E A P L W V L T R N	GCCGTGCAGGTGGCCGACGGCGAACTGGTCGATCCGGCGCAGGCCATGGTGGGGGCGGTCTC	AVQVADGELVDPAQAMVGGL	GGCCGCGTGGTCGGCATCGAGCAGCCGGGGCGCTGGGGCGGTCTGGTGGACCTGGTCGAC	GRVVGIEQPGRWGGLVDLVD	GCCGATGCCGCGTCGATCCGGTCGCTGGCCGCGGTGCTGGCGGACCCCGCGCGCG	ADAASIRSLAAVLADPRGEE	CAGGTCGCGATCCGGGCGGACGGGATCAAGGTGGCGAGGCTCGTGCCCGCCC	Q V A I R A D G I K V A R L V P A P A R	GCCGCACGCACCCGCTGGAGCCCTCGCGGCACCGTGCTGGTCACCGGCGGCACCGGAGGG	A A R T R W S P R G T V L V T G G T G G	ATCGGCGCGCACGTCGCCCGCTGGCCCGCTCGGGCGCCCGAGCACCTGGTGCTGCTG
	26443		26503		26563		26623		26683		

						72/89				-71	
26802	1	26862	1	26922	ı	26982	1	27042	1	FIG.2-71	1
++	GAHVARWLARSGAEHLVLL	GGCAGGCGCGGTGCCGACGCACCCGGCGCGTCCGAGCTGAGGGAGG	RRGADAPGASELREELTAL	GGCACGGGCGTGACCATCGCCGCCTGCGACGTCGCCGACCGGGCGCGCGC	T G V T I A A C D V A D R A R L E A V	CTCGCCGCGGAGCGCCGAGGGACGCACGGTCAGCGCCGTGATGCACGCGGGGGGTT	A A E R A E G R T V S A V M H A A G V	TCCACGTCCACGCCCCTCGACGACCTCACCGAAGCCGAGTTCACCGAGATCGCCGACGTG	TSTPLDDLTEAEFTEIADV	AAGGTGCGCGCCACCGTCAACCTGGACGAGCTCTGCCCGGACCTCGACGCGTTCGTGTTG	VRGTVNLDELCPDLDAFVL
26743 -	Н	GG 26803	ڻ و	G 26863 -	Ŋ	C 26923 -	่า	T 26983 -	S	A 27043 -	×

•						73/89			. !	FIG.2-72	
27162	ı	27222	1	27282	ı	27342	i	27402	i	FIG. 27462	ı
TTCTCCTCCAACGCGGGGGTGTGGGGCAGTCCGGGGCTCGCCTCCTACGCGGCGGCCAAC	FSSNAGSPGLASYAAN	GCCTTCCTCGACGGCTTCGCGCGCGCGCGCGGAGCGAGGGCGCGCCGGTGACGTCCATC	AFLDGFARRRSEGAPVTSI	GCCTGGGGGCTCTGGGCCGGGCAGAACATGGCCGGGGACGAGGGCGGCGGCGAGTACCTGCGC	AWGLWAGQNMAGDEGGEYLR	AGCCAGGGCCTGCGGGCCATGGACCCGGATCGGGCCGTCGAGGAACTGCACATCACCCTC	SQGLRAMDPDRAVEELHITL	GACCACGGTCAGACGTCCGTGTCGTGGACATGGATCGCAGGCGGTTCGTCGAGCTG	D H G Q T S V S D M D R R F V E L	TTCACCGCGGCCCGGCACCGGCCGCTGTTCGACGAGATCGCCGGTGCCCGGGCGGAAGCC	F T A A R H R P L F D E I A G A R A E A
27103		27163		27223		27283		Sans 27343	TITI	ST4 03	4FE7

						4703				FIG.2-7	
27522	ı	27582	ı	27642	i	27702	ı	27762	i	278	1
CGGCAGAGCGAGGAGGGCCCCGGCGCTCGCCCAGCGGCTCGCGGCGCTGTCGACGGCCGAG	RQSEEGPALAQRLAALSTAE	AGGCGCGAGCACCTCGCCCACCTGATCCGCGCCGAGGTCGCCGCGGTGCTCGGCCACGGC	R R E H L A H L I R A E V A A V L G H G	GACGACGCGGCGATCGACCGCGCGCCTTCCGCGACCTCGGCTTCGACTCCATGACC	D D A A I D R D R A F R D L G F D S M T	GCCGTCGACCTGCGGAACCGGCTCGCCGCGGTGACCGGGGTGCGGGAAGCCGCGACGGTG	A V D L R N R L A A V T G V R E A A T V	GTCTTCGACCACCCGACCATCACCCGGCTCGCCGACCACTACCTGGAGCGGCTCGTCGGC	V F D H P T I T R L A D H Y L E R L V G	GCAGCAGAGGCGGAGCAAGCCCCGGCGCTCGTGCGCGAGGTGCCGAAGGATGCCGACGAC	A A E A E Q A P A L V R E V P K D A D D
27463 -		27523	_	27583.	-	27643		27703		27763	
			6	HIDOTI	T) 177	CHEE.	т				

27	CCGATCGCGATCGTCGCCATGGCCTGCCGCTTCCCCGGCGGCGCGTGCACAACCCCGGTGAG	27882	
	PIAIVGMACRFPGGVHNPGE	i	
27	CTGTGGGAGTTCATCGTCGCCGCGGAGACGCCGTGACGGAGATGCCCACCGACCG	27942	
CHE	LWEFIVGRGDAVTEMPTDRG	ı	
N PUTITE	TGGGACCTCGACGCGCTGTTCGACCCCGACCCGCAGCGCCACGGAACCAGCTACTCGCGA	28002	
re st	WDLDALFDPOPQRHGTSYSR	75/ I	
	CACGGCGCGTTCCTCGACGGGCCGCCGACTTCGACGCGGCGTTCTTCGGGATCTCGCCG	28062	
	HGAFLDGAADFDAAFFGISP	I	
Ċ		28122	
7	D M M J) 1 1	
7	TICGAGAACGCCGGCAICGACCCGCACTCGCTGCGGGGCAGCGACACCGGCGTCTICCIC	FIG.2-74	<+
7	R E A L A M D I	2 8	.2-7

					76	5/89				1	7/2
ı	28242	i	28302	ı	28362	1	28422	ı	28482	(1	FIG. 2-/5 28542
FENAGIDPHSLRGSDTGVFL	GGCGCCGCGTACCAGGCTACGGCCAGGACGCGGTGGTGCCCGAGGACAGCGAGGGCTAC	GAAYQGYGQDAVVPEDSEGY	CIGCICACCGGCAACICCICCGCCGIGGIGICCGGCCGGGICGCCIACGIGCIGGGGCIG	L L T G N S S A V V S G R V A Y V L G L	GAAGGCCCCCGCGGTCACGGTGGACACGGCGTGTTCGTCGTCGTTGGTGGCCTTGCATTCG	EGPAVTVDTACSSSLVALHS	GCGTGTGGGTCGTTGCGTGACGGTGACTGCGGTCTTGCGGTGGCCGGTGGTGTGTCGTG	ACGSLRDGDCGLAVAGGVSV	ATGGCGGGCCCGGAGGTGTTCACCGAGTTCTCCCGCCAGGGCGGCTTGGCCGTGGACGGG	MAGPEVFTEFSROGGLAVDG	CGCTGCAAGGCGTTCTCCGCGGAGGCCGACGGCTTCGGTTTCGCCGAGGGCGTCGCGGTG
	28183		28243		28303		28363		28423	. •	28483

					77/89)			FIG.2-76	
								I		
	28602		28662		28722		28782	!	F1G 28842	
1	28	1	28	1	28	ı		1		1
>	ccgacgcccgcagggcgggrcgccaggrgcrcggcgrggrc ++++	>	AGGACGCGCGAGCAACGGTCTCGCGGCGCCCGAGCGGCGTC	>	GCAAGGCGTGGGCGCGTGCGGGGATCACCGGCGCGCGGATGTG	>	GGACCGGTACGCGGCTGGGCGATCCGGTGGAGGCGTCGGCG	æ	TTGCTGGCTACTTACGGCAAGTCGCGCGGGTCGTCGGGCCCCGGTGCTGCTGGTTCGGTG	>
æ	GTG	>	0993	Ŋ	3GAC	Ω	GTC	S	TTC	လ
>	2993	U	AGC	ഗ	200	Æ	3666	æ	366	ß
ტ	SCTC	'n	3000	Д	CACCGGCGC	Ŋ	TCCGGTGGAG	្រ	GCTGCT	IJ
េ	3GT(>	3900	Æ	CAC	H	GGT	>	GCT	7
æ	CCAC	a	AACGGTCTCGCGGCGCCC	Æ	GAT	H	TCC	Д	GGT	>
ĮΉ	ICG(ĸ	rc r	н	9 9 1	Ŋ	GTACGCGGCTGGGCGA	Ω	99000999	Ωı.
Ŋ	CGGGTCG	Ŋ	+ +	Ŋ	CGTGCGGG	æ	+ 999	Ŋ	9993	G
ſщ) 1 1 1 1	æ	CAA	z	9 9 9 9	ĸ	GCT	ч	GTC	လ
U	CAG(ĸ	GAG	လ	299 	æ	505	ĸ	GTC	ഗ
Ω	09000	æ	CGC	æ	GCGTGGGCG +	3	TAC	E	+	ტ
æ	000	Æ		Ŋ	299	æ	550	Ŋ	9093	œ
ப	CGA(Q	GGA	Ω	CAA	×	GAC	E	GTC	လ
æ	STC.	တ	CCA +	O	- - - +	8	CATGG	ប	GGCAA	×
S	STTGT	IJ	TCAACC	z	GAT	н	GCA	= ,	990	Ŋ
Įτι)) 	ĸ	GAT	н	CGT	>	AGGC 	Æ	TTA	≯
Ø	CCAC	ø	0000	A	GCG	ĸ	GGA 	ш	TAC	⊱4
노	GTCCTGCTCCAGCGGTTGT	H	GCGGCCTCGGCGATCAACC	တ	GCCCAGCAGCGCGTGATCC	ø	GCCGTGGTGGAGGCGCATG	>	TGGCTAC	Æ
υ	CCT	1	366(5 A	CCA	ø	CGT	>	GCT	H
ĸ	GT(>		A		K		A		Ħ.
	28543		28603		28663		28723		28783	

									7	<u> </u>
28902	I	28962		29022	1	29082	1	29142	(- - -	FIG.2-/
GTCGAACATCGGTCACGCGCAGGCGGCCGCGGGTGTCGCGGGCGTGATC	KSNIGHAQAAAGVAGV	CIGGGGITGAACCGCGGCCIGGIGCCGCCGAIGCICTGCCGCGGCGAGCGGICGCGGCTGCTG	LGLNPPMLCRGERSPL	ATCGAATGGTCCTCGGGTGGTGTGGAACTTGCCGAGGCCGTGAGCCCGTGGCCTCCGGCCCTCCGGCCCTCCGGCCCTCCGGCCTCCGGCCTCCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGGCCTCGTCCTCC	IE'WSSGGVELAEAVSPWAPA	GCGGACGGGGTGCGCCGGGCCGGTGTCGGCGTTCGGGGTGAGCGGGACGAACGCGCAC	ADGVRRAGVSAFGVSGTNAH	GTGATCATCGCCGAGCCCCCGGAGCCCGAGCCGCTGCCGGAACCCGGACCGGTGGGCGTG	VIIAEPEEP LPEPGPVGV	CTGGCCGCTGCGAACTCGGTGCCCGTACTGCTGTCGGCCAGGACCGAGACCGCGTTGGCA
							CIL	A		

ı	29262		. 29322	1	29382	9/89	- 29442	ı	3 - 29502	i	T FIG.2-78
, A A A N S V P V L L S A R T E T A L A	GCGCAGGCGCGCTCCTGGAGTCCGCAGTGGACGACTCGGTTCCGTTGACGGCATTGGCT	O A R L L E S A V D D S V P L T A L A	TCCGCGCTGGCCACCGGACGCCCCCACCTGCCGCGTCGTGCGGCGTTGCTGGCAGGCGAC	SALATGRAHLPRRAALLAGD	CACGAACAGCTCCGCGGCAGTTGCGAGCGGTCGCCGAGGGCGTTGCGGCTCCCGGTGCC	HEQLRROOLRAVAEGVAAPGA	ACCACCGGAACCGCCTCCGCCGGCGCGTGGTTTTCGTCTTCCCAGGTCAGGGGGTGCTCAG	T'T G T A S A G G V V F V F P G Q G A Q	TGGGAGGGCATGGCCCGGGGCTTGCTCTCGGTCCCCGTCTTCGCCGAGTCGATCGCCGAG	WEGMARGLLSVPVFAESIAE	TGCGATGCGGTGTTGTCGGAGGTGGCCGGGTTCTCGGCCTCCGAAGTGCTGGAGCAGCGT
H	G 29203 -	æ	T 29263 -	S	C 29323 -	Ħ	29383 -	C.	129443 -	,	ζ,

CCGGACGCCGCCGTCGAGCGGGTCGACGTCGTACAGCCGGTGTTGTTCTCCGTGATG	Q P V L F S V M -	GTGTCGCTGGCGCGCTGTGGGGCGCTTGCGGAGTCAGCCCCTCGGCCGTCATCGGCCAT ++	SPSAVIGH-	TCGCAGGGCGAGATCGCCGCGCGGGGTGGTGGCCGGGGTGTTGTCGCTGGAGGACGGCGTG	V L S L E D G V -	GCGGGCAAGGGCGGCATGGTC 29802	LAGKGGMV -	TCGTTGGCGGCTCCCGGTGAACGCGCCCCGCGCGCTGATCGCACCGTGGGAGGACCGGATC ++	IAPWEDRI - FIC 7-70	
+	α >	AGTCAG	s N	CGGGGT(ტ	STGCGCTG	A	CGCTGAT	H	#COGTOGECOGGTCAACTCCCCGGTCGTCGTGGTCTCCGGCGATCCGGAGGCGCTG
	D V	TTGCGC	_ຽ	TGGTGG(+	V	GCGTTGC	J R	+	R	CTCGG
3CGGGT	я >	366660	G B		A A	GAAGGC	K A	ACGCG	R A	
CTGGAC	<u>а</u> 1	CTGTG	L W	+	A	+	R A	CGGTGA	<u>ម</u>	i (
CGTCG	S S	 3CGCGG	A	SAGATC	H B	SCCTGCG	A L	GCTCC	A P	(
ACGCGCC	Æ	CGCTGG	н	TCGCAGGGCGAGA	ტ	TCGTGGC	>	TCGTTGGCGGCTC	æ	
	PD		> S		o s		R		S	
29563		29623		29683		29743		29803		

						82/	'8 9		Ċ	∞	
30282	1	30342	i	30402	1 2 3 3 5 1	30462	1	30522	Ī	FIG. Z-8 30582	ĵ
i	0.	AC		AC.	_	AC	0	3.GC	(D	AGGC	Ŋ
+	w .	CGTGCAC	Ħ	CCCAAC	Z	CCGAC	Ω	+	G	GAC	
	G	ACG.	>	IGC		FCG.	K	AAG 	U	UCG	Ŋ
	Н	, GC2	H	GGTGGCGCT	Н	AGCT	្ន	rgg.	ា	AGG	K
. .	K	+	æ	999	. 4	GACCA	ø	SGACCT	H	GAGAJ	×
 	>	900	ø	GGT	>	CGA	Ω	GGA	Ω	CGP	ា
 	K	292	K	JUL I	ىم	GTC	လ	GGT 	>	AGT	>
į	Ω	JC T	h	ACC.	വ	3GT	>) 	Δι) - 	4
†	· 4	CGA(+-	ធ	GGCACC	4	GGAG +-	ជ	CACG	H	CAG(လ
İ	Æ	000	æ	3,000	4))) (ш	CAC	H	SAC	H
	н	ATC	Ĥ	SC	بم	3600	Ø	3600	Æ	GCT	ı
+	凹	CCTC	H	TCTTCC	- [<u>t</u>	GCTC	卢	GCT(-+-	ı	GTC(-+-	တ
	a	CAC	Ħ	GTC	>	TGC	Z	JCCC	Д	3GA(ធ
	>	GAG	ជ	AAC	z	TAC	×	3CG1	~	VCCC	Δι
	Æ	GGAGGAGCACCTGATCGCCGAGCTCGCCCGGGCGCACGTGCAC	ជ	GCGGAACGTCTTCCCGGCGCACCTCCGGTGGCGCTGCCCAAC		GCGGTACTGGCTCGCCGGAGGTGTCCGACCAGCTCGCCGAC	ĸ	CTGGCGACCGCTGGCCACCACGCCGGTGGACCTGGAAGGCGGC	3	CGC7	æ
†	4 .	5000	æ) TGG	3	SCAC	a	CGAC	Ω	TTCCTGGTCCACGGGTCCGCACCGGAGTCGCTGACCAGCGCAGTCGAGAAGGCCGGAGGGC	တ
	H	CTGCACCGCGACACCGC	H	GGCGTGGCCGTGGACTG	Ω	TACCCGTTCGAGCCCCA	Д	AGCCGCTACCGCGTCGA	>	0000	Ŋ
į	н	CGA(Ω	CGT	>	CGA(ធ)) 	ĸ	CCA	Ħ
+	>	+	¤	Greecer +	4	TTCG	Įzi	CCGCTAC	×	GGTC +-	>
 	Д	CAC	Ħ	GTG	>	CCC	ь	990	K	CTC	ы
- 		CTG	H	095 111	ტ	TAC	≯	AGC	S	TTC	ŢŦ
30223		30283		30343)	30403		30463		30523	

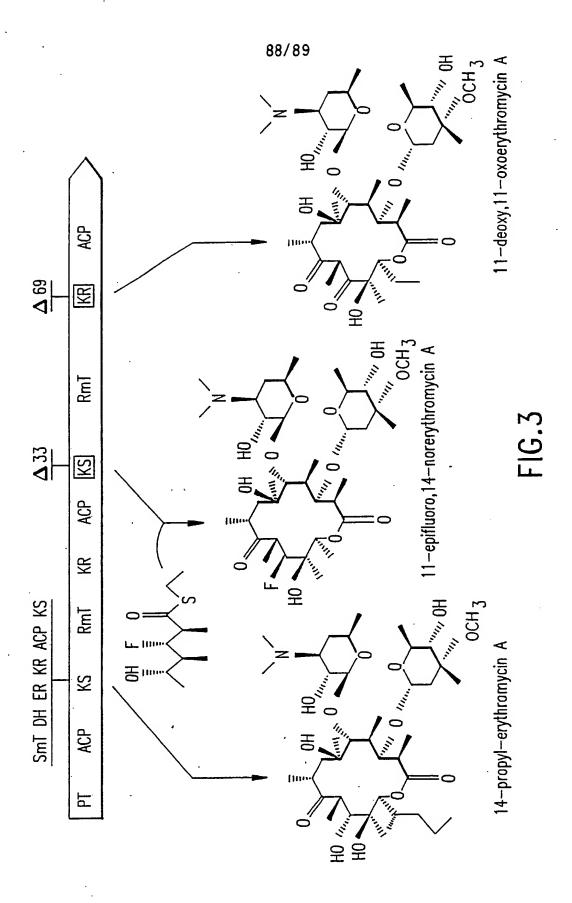
						83/89)		C	F10.2 - 02 1942	
30642	1	30702	í	30762	ı	30822	ı	30882	ر ا ا	FIG. 30942	i
CGCGTCGTGCCGGTCGCCTCGGCCGACCGCGAAGCCTCGGCGGCCCTGCGGGAGGTGCCG	RVVPVASADREASAALREVP	GGCGAGGTCGCCGGCGTGCTCTCGGTCCACACCGGCGCCGCAACGCACCTCGCCCTGCAC	GEVAGVLSVHTGAATHLALH	CAGTCGCTGGGTGAGGCCGGCGTGCGGCCCCGCTCTGGCTGG	Q S L G E A G V R A P L W L V T S R A V	GCGCTCGGGGAGTCCGAGCCGATCCCGAGCAGGCGATGGTGTGGGGGTCTCGGGCGCGC	ALGESEPVDPEQAMVWGLGR	GTCATGGGCCTGGAGACCCCGGAACGGTGGGGCGGTCTGGTGGACCTGCCCGCCGAACCC	V M G L E T P E R W G G L V D L P A E P	GCGCCGGGGGACGGCGAGGCGTTCGTCGCCTGCCTCGGCGCGGGCGG	APGDGEAFVACLGADGHEDQ
30583		30643		30703		30763		30823		30883	

					8	84/89			FIG 2,83	7
31002	1	31062	į	31122	ı	31182	ı	31242	_ 	'n
GICGCGAICCGIGACCACGCCCGCIACGGCCGCCGCCTCGICCGCGCCCCCGCTGGGCACC 3+++++++-	VAIRDHARYGRRLVRAPLGT	CGCGAGTCGAGCTGGGAGCCGGCGGCACGGCGCTGGTCACCGGCGCGCACCGGTGCGCTC	RESSWEPAGTALVTGGAL	GGCGGCCACGTCGCCCGCCACCTCGCCAGGTGCGGGGTGGAGGACCTGGTGCTGGTCAGC	G G H V A R H L A R C G V E D L V L V S	AGGCGCGGCGTCGACGCTCCCGGCGGCCGAGCTGGAAGCCGAACTGGTCGCCCTCGGC	RRGVDAPGAAELEAELVALG	GCGAAGACGACCATCACCGCCTGCGACGTGGCCGACCGCGAGCAGCTCTCCAAGCTGCTG	AKTTITACDVADREQLSKLL	GAAGAACTGCGGGGAGGGACGTCCGGTGCGGACCGTGCACACCGCCGCCGGGGTGCCC 13+++++++
30943		31003		31063		31123		31183		31243

						85 / 8	9				FIG. Z - 84 662
ı	31362	ı	31422	1	31482	ı	31542	ı	31602	ı	710 31662
E E L R G Q G R P V R T V V H T A G V P	GAATCGAGGCCGCTGCACGAGATCGGCGAGCTGGAGTCGGTCTGCGCGGGGGAAGGTGACC	ESRPLHEIGELESVCAAKVT	GGGGCCCGGCTGCTCGACGAGCTGTGCCCGGACGCCGAGACCTTCGTCCTGTTCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCGTCCTCTCGTCCTCC	GARLLDELCPDAETFVLFSS	GGAGCGGGGGTGTGGGGCAGTGCGAACCTCGGCGCCTACTCCGCGGCCAACGCCTACTC +++++++-	G A G V W G S A N L G A Y S A A N A Y L	GACGCGCTGGCCCACCGCCGCGGTGCGGAAGGCCGTGCGGCGACGTCCGTC	DALAHRRAEGRAATSVAWG	GCCTGGGCGGGGGAGGGCATGGCCACCGGCGACCTCGAGGGGGCTCACCCGGCGCGCGC	AWAGEGMATGDLEGLTRRGL	CGCCCGATGGCGCCCCGAGCGCGCGCGCGCGCTGCACCAGGCGCTGGACAACGGCGAC
	31303		31363		31423		31483		31543		31603

						86/8	9			0	0
	. •									Ć	7
	2		82		42		02		62		F10. Z = 322
	31722	1	31782	1	31842	ı	31902	i	31962	ı	32022
•		•		•		·			8		AC
Ω	+	æ	CGGTG	>	3GCAC	I	rcacc +	. H	ccacc +	H	CCGAC
O		Æ	0500	Ø	ACTO	တ	CGT	Ĺı	AGG(æ	TGG III
Z	TCAC	H	000	Ω	3GC/	H	4GCC	<u>Ω</u> ,	AGC.	Ø	GGT
Д	CGACGICGACIGGGAGGCCIICGCGGICGGCIICACCGCCGCC +++	Įτι	GGACGAGCTCGTCACGCCGGCGGTGGGGGGCCGTCCCCCGCGGTG +++++++	>	CAGGCGCCCCGGCGCGGGAGATGACGTCGCAGGAGTTGCTGGAGTTCACGCACTCGCAC +++++++	H	GCATICCAGCCCGGACGCGGTCGGGCAGGACCAGCCGTICACC +++	Ø	GCTGACCGCGGTCGGGCTGCGCAACCAGCTCCAGCAGGCCACC +++++++	Ø	GGGCTCGCGCTGCCCGCGACCCTGGTGTTCGAGCACCCCCACGGTCCGCAGGTTGGCCGAC +++++++-
H	CGGCT	ტ	0999	Æ	\GTJ	Ĺų	4GG7	Ω	AGC	1	HCC I
K	GGT	>	9990	Ŋ	GTTGCTGGA +	ជា	3GCA	Ø	ACC	Ø	555
a	090	Æ		>	rGC1	н	GGTCGG	ტ	CGCA.	z	CCA +
H	CTTCG	ſτι	0993	æ	GT7	н	3663	>	IGC(X	CACCCC
н	299	A.	009:	Д		딦	ACGC	Ø	3GC	H	AGC.
Ø	GGA	田	CAC	H	(C)	ø	2667	Ω	rcg(Ŋ	TGTTCGAG+
ĸ	CTGGG -+	3	CGT	>	ACGTC	လ	3000	Д	1997 -+	>	HGT +
н	CGA	Ω	GCT	1	GAC	H	CAG	ဟ	2000	K	rgg
Ø	CGT	>	CGA	<u></u> ы	GAT	Σ	TTC	က	rgac	EH	30C
ĸ	CGA +	Q	GGA ++	Ω	+-	ជ	3GC.P	x	7,6C3 ++	H	CGAC
凹	ACGTGCGTTTCGATCGC	Æ	CGGCCGCGTCCGCTGCT	н	929	ĸ	GTCGCGGCGATCCTCGG	Ŋ	GAGCTCGGCTTCGACTC	လ))
Д	GAT	H	GCT	H	299	K	CCT	Н	,cg	Q	
æ	FTC 	လ	1CC	Д	CAGGCGGCCCC	Д	GGCGAT	Н	CTI	ĮΞij	CGCGCTGC
Σ	GCGTT	>	GCGT	ĸ	+	Æ	966	A	CGGCT	Ŋ	090
Q,	STG(ပ	Ü U U U	Д	286	æ	292	K	GAGCT	н	GCI
ĸ	ACC	Ħ		~		ø		>		떠	
	31663		31723		31783	•	31843		31903		31963

					87/	'8 9			FIG.2-86	
	32082		32142		32202	:	32262		F)(
1		ı		ı				1		·
Ω	CAGCGGGACTCCCGCCCGGGAAGCGAGCAGCGCTCTTCGC	ĸ	GTCGGGCAGGGTCCGACCTCGACCTGCTGGCG	Æ	TTCGACGGCTCCGACGGGTTCTCCCTCGATCTCGTG	>	GTCACGGTGATCTGCTGCGCGGGAACGGCGGCGATC	H	GGCGCTGCGCGGAATCGCTCCGGTTCGG	æ
Ø	TCT	н.	GCT	Н	TCI	ъ.	2660	Æ	2663	>
H	292	æ	CCI	ы	,CG2	Ω)992	æ	CTCC	ር <mark>ተ</mark>
ĸ	CAG +	လ	CGA	Ω) +-	. н	SAAC	EH	rcg(A
Ж	GAG	လ	CCI	H	CIC	ഗ)993	Ŋ	SAA	H
>	AGC	æ	CTA	X	GTI	Ĺų	3000	K	3000	Ŋ
H	GGA	ធ	GTC	လ	090	Ŋ	SCTC	U	rGCC	ፎ
Ω	500	K	000	ц	100 A	Ω	CTC	U		H
= ,	292	æ		>	CIC	S	GAJ	Н	3660	K
ы	HCC 	Δ,	CAG	ద		Ŋ	3661	> .	CTCGCCGG	ပ
Įτί	GAC	H	-+-	ပ	-+-	Ω	CAC	H	, +	A
>	990	ტ	GTC	ഗ	CTI	Įщ	\GG1	>	3GCJ	H
н	CAG	လ	CGT	>	GCA	H	AGP:	ជ	SS	ĸ
H	CGA +-+	Ω	+	Ŋ	CGA	ជា	+-	Ŋ	TTCAC	H
æ	GCT	ч	299	æ	900	ĸ	TGGCCGACGGTCCCGG		GTJ	Ĺ
Ω	GCA	G Q Q L	GCA	я 0	CTI	Ĺτι	0001	ტ	CG.	ជា
н	ACA	ø	500	ĸ	GGA	S D	CGA	A D G	360.2	Ħ
æ	AGG +	Ŋ	CTA 	×	GTC 		+ +	A	1100	Д
T	CACATAGGACAGCAGCTCGA	н	GACGGCTACCGGCAGGCGGGGGGGTCGGGGTCCGGTCCTACCTCGACCTGCTGGCG	ტ	GGGCTGTCGGACTTCCGCGACTTCGACGGCTCCGACGGGTTCTCCCTCGATCTCGTG	ų.	GACATGGCCGACGGTCCCGGAGAGGTCACGGTGATCTGCTGCGCGGGAACGGCGGCGATC	Σ	TCCGGTCCGCACGAGTTCACCCGGCTCGCCGGGGCGCTGCGCGGAATCGCTCCGGTTCGG	Ŋ
Ü	CA	Ħ		Ω		ტ		Ω		ິທີ
	32023		32083		32143		32203		32263	



SUBSTITUTE SHEET

NUMBER	SITE	DISTANCE (Kb)a
1	BamHl	-3.60
2 -	Pvull	-3.50
-3	Pvull	-3.40
4	PstI	-3.05
5	BamHl	-2.95
6	Xhol	-2.80
7	PstI	-2.00
8	Hindll	-1.60
9	Sphl	-1.55
10	EcoRI	-1.50
11	· Kpnl	-1.35
12	EcoRI	-1.05
13	Smal ^b	-0.90
14	Sphl	-0.75
15	Kpni	-0.65
16	Smal	-0.20
10	Jillui	

FIG.4

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/00427

	The state of the s						
1. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) ³							
According to International Patent Classification (IPC) or to both National Classification and IPC							
IPC (5): A01N 43/22; A61K 31/71; C07H 17/08; C12N 1/21; C12P 19/62 US CL : 435/76, 252.35, 886; 514/29; 536/7.2							
II. FIEL	DS SEAR		and a Constant				
	Minimum Documentation Searched 4						
Classificat	on System		Classification Symbols				
U.S. 435/76, 183, 252.35, 88		435/76, 183, 252.35, 88	6; 514/29; 536/7.2				
Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched 5							
APS, BIOSIS, MEDLINE, BIOTECH. ABSTRACTS							
III. DOC	UMENTS	CONSIDERED TO BE RELEVANT 14					
Category*		n of Document,18 with indication, where app	propriate, of the relevant passages ¹⁷	Relevant to Claim No. 18			
У	J. of 1985, Erythi pages	1-30					
У	J. of J.M. Erythi 2372-2	1-30					
У	US, A, 4,874,748 (Katz et al) 17 October 1989, see column 1, lines 41-59; column 3, lines 47-60.			1-30			
У .	US, A, 4,935,340 (Baltz et al) 19 June 1990, see entire 1-29 document.						
×							
* Special	notecnica	of cited documents: 15	"T" later document published afte	the international filing			
"A" downward defining the general state of the set which is							
not	not considered to be of particular relevance application but cited to understand the principle or						
inter "L" doce	arlier document but published on or after the ternational filing date "X" document of particular relinvention cannot be consider considered to involve an invention.			ed novel or cannot be			
ano	or which is cited to establish the publication date or shorter citation or other special reason (as specified) "Y" document of particular relevance; the claim invention cannot be considered to involve						
or other means "P" document published prior to the international filing date but later than the priority date claimed inventive step when the document is combing one or more other such documents, such combing obvious to a person skilled in the art document member of the same patent family							
IV. CERTIFICATION							
Date of the Actual Completion of the International Search ² 20 MARCH 1992			Date of Mailing of this International 3 0 MAR 1992	Search Report ²			
International Searching Authority ¹			Signature of Authorized Officer 20	Maril 1			
ISA/US			Dian Cook	Name for			